



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 157801

TO: Celine Qian
Location: REM/2A64/2C70
Art Unit: 1636
Tuesday, July 05, 2005

Case Serial Number: 09/936271

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Celine Qian Examiner #: 78710 Date: 6/28/05
 Art Unit: 1636 Phone Number 302-0777 Serial Number: 09/936271
 Mail Box and Bldg/Room Location: 2A 64 Results Format Preferred (circle): APER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of invention: Novel Human Kallikrein-like genes

Inventors (please provide full names): Yousef et al.

Earliest Priority Filing Date: 9/10/2007

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

search SEQ ID NO: 13 : 1 - 5,000 bp

NA 11570

6000 - 7000 bp

8000 - 9000 bp

10000 - 11,570 bp

search SEQ ID NO: 14

AA 293

both commercial database & interference.

ME

STAFF USE ONLY

Type of Search		Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>4</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) <u>1</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>7/1/05</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>7/5/05</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep. Review Time _____	Fulltext _____	Sequence Systems <u>DBA</u> <u>DBA</u>
Clerical Prep. time: _____	Patent Family _____	WWW/Internet _____
Online Time _____	Other _____	Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 18:09:16 ; Search time 4527.92 Seconds

(without alignments)
16811.954 Million cell updates/sec

Title: US-09-936-271c-13_COPY_10000_11570

Perfect score: 1571

Sequence: 1 agggaggtatgggaattga.....cccgagaataaactgagaag 1571

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1571	100.0	11570	6	CQ788219	Sequence
2	1571	100.0	11570	6	CQ874885	Sequence
3	1571	100.0	11570	6	CQ874961	Sequence
4	1571	100.0	11570	9	AF135028	Homo sapi
5	1571	100.0	107487	9	AC011483	Homo sapi
6	1571	100.0	230000	9	AF243527	Homo sapi
7	1555.2	99.0	217346	2	AC027602	Homo sapi
8	1482	94.3	200792	2	AC130782	Pan trogl
9	482	30.7	586	6	AX331407	Sequence
10	482	30.7	586	6	AX331824	Sequence
11	482	30.7	586	11	G41906	SHGC-56840
12	482	30.7	735	6	AX429955	Sequence
13	482	30.7	738	6	BD139877	Compounds
14	482	30.7	1143	6	CQ720600	Sequence
15	482	30.7	1301	9	AY461805	Homo sapi
16	482	30.7	1370	9	AY279380	Homo sapi
17	482	30.7	1381	6	BD107879	Homo sapi
18	482	30.7	1387	6	CQ874884	Sequence
19	482	30.7	1387	6	CQ874960	Sequence

20	482	30.7	1387	9	AF168768	Homo sapi
21	482	30.7	1438	9	AY279381	Homo sapi
22	482	30.7	1439	6	BD107880	36 human
23	482	30.7	1499	6	AR352504	Sequence
24	482	30.7	1499	6	BD005362	Protease.
25	482	30.7	1516	6	BD107865	36 human
26	482	30.7	1527	9	BC008036	Homo sapi
27	482	30.7	1570	6	AR252580	Sequence
28	482	30.7	1570	6	AR528679	Sequence
29	482	30.7	1570	6	AX080829	Sequence
30	482	30.7	1570	6	AX403421	Sequence
31	482	30.7	1570	6	AX464322	Sequence
32	482	30.7	1570	9	AY359010	Homo sapi
33	476	30.3	1476	6	AR078184	Sequence
34	476	30.3	1476	6	AR137506	Sequence
35	476	30.3	1476	6	AR242358	Sequence
36	476	30.3	1476	6	BD082659	Keratinoc
37	460	29.3	1504	6	AR263823	Sequence
38	449	28.6	673	6	AX067353	Sequence
39	315.4	20.1	498	11	BV209338	KLK5_3564
40	236.6	15.1	321519	2	AL714004	Homo sapi
41	236	15.0	175559	9	AC112498	Homo sapi
42	234.4	14.9	174380	2	AC016075	Homo sapi
43	233.2	14.8	151880	9	AP005716	Homo sapi
44	231.2	14.7	79516	9	AC004834	Homo sapi
45	231.2	14.7	156949	9	AL161652	Human DNA

ALIGNMENTS

RESULT 1

CQ788219

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

location/Qualifiers

1. 11570

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match	100.0%;	Score 1571;	DB 6;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1571;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
Qy	1	AGGGAGGTATGGGAATTGAAGACAGGAAACACAAATTAGTCCAAGCGAATGGATTCTA	60	
Db	10000	AGGGAGGTATGGGAATTGAAGACAGGAAACACAAATTAGTCCAAGCGAATGGATTCTA	10059	
Qy	61	TTGGGAGTGAATTCGCCCCCTAGGAAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA	120	
Db	10060	TTGGGAGTGAATTCGCCCCCTAGGAAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA	10119	
Qy	121	CACCTATATGAGGGGCAATTTACTGGCAACTAATGATAGATGCAAGTGTCTGTTCAAC	180	
Db	10120	CAACTATATGAGGGGCAATTTACTGGCAACTAATGATAGATGCAAGTGTCTGTTCAAC	10179	
Qy	181	ATGCTATGATGCACAGCGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC	240	
Db	10180	ATGCTATGATGCACAGCGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC	10239	

QY	241	AGTCCAGATCGAGGAACCTTCATCCAGGGCTTGAAACCGGTATTTTTCAGAGGGAG	300
Db	10240	AGTGGCCAGATCGAGGAACCTTCATCCAGGGCTTGAAACCGGTATTTTTCAGAGGGAG	10299
QY	301	GTATAAGGATGGGTGGTGAGAAATGGGGAAGAGGTGTGTCCAGTAAGAGAAATAA	360
Db	10300	GTATAAGGATGGGTGGTGAGAAATGGGGAAGAGGTGTGTCCAGTAAGAGAAATAA	10359
QY	361	GGCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGGAGCGGAGAGATACACGATGA	420
Db	10360	GGCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGGAGCGGAGAGATACACGATGA	10419
QY	421	GGGAGACAGGCTGGAAACAGAAAGTAGAGACCAAGATTCGAGATGTGGAGAGGAAGGTCA	480
Db	10420	GGGAGACAGGCTGGAAACAGAAAGTAGAGACCAAGATTCGAGATGTGGAGAGGAAGGTCA	10479
QY	481	CAGACCCCCCGAAATGATGTGGACAAACAGGAATCTGGAAGAGGAAGATGGAGTGGAG	540
Db	10480	CAGACCCCCCGAAATGATGTGGACAAACAGGAATCTGGAAGAGGAAGATGGAGTGGAG	10539
QY	541	AGTGACAAATGGGGTCTAAAGTTGAACCTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA	600
Db	10540	AGTGACAAATGGGGTCTAAAGTTGAACCTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA	10599
QY	601	ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTGAGGCCAGGAGTTCGAGACCAG	660
Db	10600	ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTGAGGCCAGGAGTTCGAGACCAG	10659
QY	661	CTGGGCCAATGGTGAAACCCCGTCTCTACAAAAAATAATACAAAAAATPAGCCGGGTGT	720
Db	10660	CTGGGCCAATGGTGAAACCCCGTCTCTACAAAAAATAATACAAAAAATPAGCCGGGTGT	10719
QY	721	GGTGATGGACACCTGTAGTCACAGCTACTTGGAGGCTGAGGACGAGGATTCCTTGAAC	780
Db	10720	GGTGATGGACACCTGTAGTCACAGCTACTTGGAGGCTGAGGACGAGGATTCCTTGAAC	10779
QY	781	CCGGGAGATGGAGGCTGCAGTGAGCTGAGTTCAGGCCACTTGCCTCCAACTCGGCAACA	840
Db	10780	CCGGGAGATGGAGGCTGCAGTGAGCTGAGTTCAGGCCACTTGCCTCCAACTCGGCAACA	10839
QY	841	GAGTAAGACTCCATCTCAAAAAAATAAGCTGGATTTGGAGTGAATATTAATAACATT	900
Db	10840	GAGTAAGACTCCATCTCAAAAAAATAAGCTGGATTTGGAGTGAATATTAATAACATT	10899
QY	901	CTCCCTCTCTCTTTTGGCTGTGTCTCCATCTGTCTTTTTCGCAATTTCTCATCT	960
Db	10900	CTCCCTCTCTCTTTTGGCTGTGTCTCCATCTGTCTTTTTCGCAATTTCTCATCT	10959
QY	961	CTGTACTTTCCATCTCTGTGTGTCTGTCCCATCTGCTTCCCATCTATGGGCATCTCTG	1020
Db	10960	CTGTACTTTCCATCTCTGTGTGTCTGTCCCATCTGCTTCCCATCTATGGGCATCTCTG	11019
QY	1021	GGTCTCTCATGTCTCTTCGCCACATTTGGCCACATCTCTGCTCTCTCATGCCCCCTTT	1080
Db	11020	GGTCTCTCATGTCTCTTCGCCACATTTGGCCACATCTCTGCTCTCTCATGCCCCCTTT	11079
QY	1081	TCTCTCTCGAGGTGATTTCTGGGGGCGCTGTGGTCTGCATAGGCTCCCTCGAGGACATC	1140
Db	11080	TCTCTCTCGAGGTGATTTCTGGGGGCGCTGTGGTCTGCATAGGCTCCCTCGAGGACATC	11139
QY	1141	GTGTCTCTGGGAGATTACCTTGTGGCCGCCCAACAGACCGGGTGTCTACAGAACCTTC	1200
Db	11140	GTGTCTCTGGGAGATTACCTTGTGGCCGCCCAACAGACCGGGTGTCTACAGAACCTTC	11199
QY	1201	TGCAAGTTCACCAAGTGGATTCACGAAACCATCCAGGCCAACTCTCTGAGTCAATCCAGGA	1260
Db	11200	TGCAAGTTCACCAAGTGGATTCACGAAACCATCCAGGCCAACTCTCTGAGTCAATCCAGGA	11259
QY	1261	CTCAGCACACCGGCATCCCCACTGTCTGAGGAGACAGCCCTGACATCTCTTTCAGACCTT	1320
Db	11260	CTCAGCACACCGGCATCCCCACTGTCTGAGGAGACAGCCCTGACATCTCTTTCAGACCTT	11319
QY	1321	CATTCTCTCCAGAGATGTTGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCT	1380
Db	11320	CATTCTCTCCAGAGATGTTGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCT	11379
QY	1381	GGACTCAGGGTCTGCTTCCCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTTGAACCCCTGGG	1440
Db	11380	GGACTCAGGGTCTGCTTCCCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTTGAACCCCTGGG	11439
QY	1441	ACAAATTTCCAAACCTGTCCAGGGCGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCA	1500
Db	11440	ACAAATTTCCAAACCTGTCCAGGGCGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCA	11499
QY	1501	TCCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAAGAAAT	1560
Db	11500	TCCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAAGAAAT	11559
QY	1561	AACTGAGAAG 1571	
Db	11560	AACTGAGAAG 11570	
RESULT 2			
CQ874885			
LOCUS			
DEFINITION			
Sequence 3 from Patent WO2004077060.			
ACCESSION			
CQ874885			
VERSION			
CQ874885.1 GI:52748035			
KEYWORDS			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Dianandis, E.P. and Petraki, C.			
Assay for detection of renal cell carcinoma			
Patent: WO 2004077060-A 3 10-SEP-2004;			
Mount Sinai Hospital (CA)			
FEATURES			
source			
1..11570			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity 100.0%; Score 1571; DB 6; Length 11570;			
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	AGGGAGGTATGGGAATTTGAAGACAGCAAGAAACACAAATTTAGTCCAAAGCAATGATTCTTA	60
Db	10000	AGGGAGGTATGGGAATTTGAAGACAGCAAGAAACACAAATTTAGTCCAAAGCAATGATTCTTA	10059
QY	61	TTGGGAGTGAATTTGCCCCCTAGAAGACACTGGCAATCCAGGAGACATTTTGGTTGTCA	120
Db	10060	TTGGGAGTGAATTTGCCCCCTAGAAGACACTGGCAATCCAGGAGACATTTTGGTTGTCA	10119
QY	121	CAACTATATGGAGGGGCAATTTCTGCAACTAATGGATAGATGCCAAGTGTGCTTTCAAC	180
Db	10120	CAACTATATGGAGGGGCAATTTCTGCAACTAATGGATAGATGCCAAGTGTGCTTTCAAC	10179
QY	181	ATGCTATGATGCACACGGCAGGCTCCACAAACAAACATTTATCCAGCTTCAGATGCCAC	240
Db	10180	ATGCTATGATGCACACGGCAGGCTCCACAAACAAACATTTATCCAGCTTCAGATGCCAC	10239
QY	241	AGTGCCAGATCGAGGAACCTTCATCCAGGGGCTGAGAACCGGTATTTTTCAGAGGGAG	300
Db	10240	AGTGCCAGATCGAGGAACCTTCATCCAGGGGCTGAGAACCGGTATTTTTCAGAGGGAG	10299
QY	301	GTATAAGGATGGGTGGTGGAGATGGGGAAGAGTGTGTCTCAGTAAGAGAAATAA	360
Db	10300	GTATAAGGATGGGTGGTGGAGATGGGGAAGAGTGTGTCTCAGTAAGAGAAATAA	10359
QY	361	GGCCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGGAGCGGAGAGATACACGATGA	420

Db 10360 GGCCTGCACAGGCTGGAGGGAGAGTCAGAGAGAAAGGGAGCGAGAGATACAGATGA 10419
QY 421 GGGAGACAGGCTGGAAACAGAAAGTAGAGACAGAGATTTCGAGATGTGGAGAGGAAGGTCA 480
Db 10420 GGGAGACAGGCTGGAAACAGAAAGTAGAGACAGAGATTTCGAGATGTGGAGAGGAAGGTCA 10479
QY 481 CAGACCCCGGAAATGATGTGGCAACACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 540
Db 10480 CAGACCCCGGAAATGATGTGGCAACACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 10539
QY 541 AGTCACAAATGGGCTTAAAGGTTGAATTTGGAGGCCAGGATGGTGGCTCAGGCCCTGTA 600
Db 10540 AGTGACAAATGGGCTTAAAGGTTGAATTTGGAGGCCAGGATGGTGGCTCAGGCCCTGTA 10599
QY 601 ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTGGAGGCCAGGATTCGAGACCAG 660
Db 10600 ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTGGAGGCCAGGATTCGAGACCAG 10659
QY 661 CTTGGCCCAACATGGTGAACCCCGTCTCTACAAAAAATACAAAAATTTAGCGGGTGT 720
Db 10660 CTTGGCCCAACATGGTGAACCCCGTCTCTACAAAAAATACAAAAATTTAGCGGGTGT 10719
QY 721 GGTGATGACACCTGTAGTCACAGCTACTTTGGAGGCTGAGGAGGAGATTTGCTTCAAC 780
Db 10720 GGTGATGACACCTGTAGTCACAGCTACTTTGGAGGCTGAGGAGGAGATTTGCTTCAAC 10779
QY 781 CCGGGAGATGAGGCTGACGTAGCTGAGTCAAGGCACTGCGCTCCAACTCGGGCAACA 840
Db 10780 CCGGGAGATGAGGCTGACGTAGCTGAGTCAAGGCACTGCGCTCCAACTCGGGCAACA 10839
QY 841 GAGTAGACTCCATCTCAAAAAAAGCTGAGTGGATGAGTGAATATTAATTAACATT 900
Db 10840 GAGTAGACTCCATCTCAAAAAAAGCTGAGTGGATGAGTGAATATTAATTAACATT 10899
QY 901 CTCCCTCTCTCTCTTTGCTGTCTGCTCCTATCTCTCTTTTCTCAATTTCTTCACT 960
Db 10900 CTCCCTCTCTCTCTTTGCTGTCTGCTCCTATCTCTCTTTTCTCAATTTCTTCACT 10959
QY 961 CTGTACTTCCATCTCTGTGTCTGTCTCCATCTGCTTCTCCATCTATGGCATCTCTG 1020
Db 10960 CTGTACTTCCATCTCTGTGTCTGTCTCCATCTGCTTCTCCATCTATGGCATCTCTG 11019
QY 1021 GGTCTCTCATGTCTCTTCTGCCCCACATTTGGCAATCTCTGCTCTCTCATGCCCCCTT 1080
Db 11020 GGTCTCTCATGTCTCTTCTGCCCCACATTTGGCAATCTCTGCTCTCTCATGCCCCCTT 11079
QY 1081 TCTCTCTGAGGCTGATTTCTGGGGGCTGTGTCTGCAATGGCTCCCTGCGAGGACTC 1140
Db 11080 TCTCTCTGAGGCTGATTTCTGGGGGCTGTGTCTGCAATGGCTCCCTGCGAGGACTC 11139
QY 1141 GTGTCTCTGGGAGATTACCTTTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTC 1200
Db 11140 GTGTCTCTGGGAGATTACCTTTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTC 11199
QY 1201 TGCAGTTCACCAAGTGGATTCAGAAACCATCCAGGCCCAACTCTCGAGTCACTCCAGGA 1260
Db 11200 TGCAGTTCACCAAGTGGATTCAGAAACCATCCAGGCCCAACTCTCGAGTCACTCCAGGA 11259
QY 1261 CTCAGCACCGGATCCACCTGCTGCGGGGACGCTGACACTTCTTTCAGACCTT 1320
Db 11260 CTCAGCACCGGATCCACCTGCTGCGGGGACGCTGACACTTCTTTCAGACCTT 11319
QY 1321 CATTCCTTCCAGAGATTGAGAAATGTTATCTCTCCAGGCCCTGACCCCATGCTCTCT 1380
Db 11320 CATTCCTTCCAGAGATTGAGAAATGTTATCTCTCCAGGCCCTGACCCCATGCTCTCT 11379
QY 1381 GGACTCAGGCTGCTCTTCCCCCACTTTGGGCTGACCGTGTCTCTTAGTGAACCTCTGG 1440
Db 11380 GGACTCAGGCTGCTCTTCCCCCACTTTGGGCTGACCGTGTCTCTTAGTGAACCTCTGG 11439
QY 1441 AACAAATTCCTCCAGGCGGGGTTGGCTCTCAATCTCCCTTGGGGCACTTCA 1500

Db 11440 AACAAATTCCTCCAGGCGGGGTTGGTCTCAATCTCCCTGGGCACTTCA 11499
QY 1501 TCCTCAAGCTCAGGGCCATCTCTTCTGACGCTTGACCCAAATTTAGTCCAGAAAT 1560
Db 11500 TCCTCAAGCTCAGGGCCATCTCTTCTGACGCTTGACCCAAATTTAGTCCAGAAAT 11559
QY 1561 AAACCTGAGAAG 1571
Db 11560 AAACCTGAGAAG 11570
RESULT 3
LOCUS CO874961 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION CO874961
VERSION CO874961.1 GI:52748060
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Diamandis, E.P.
TITLE Multiple marker assay for detection of ovarian cancer
JOURNAL Patent: WO 2004075713-A 5 10-SEP-2004;
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
Location/Qualifiers
1. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1571; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGAGGATATGGGAATTGAAGACAGAAACACAAATTAAGTCCAAAGCAATGGATTCTA 60
Db 10000 AGGAGGATATGGGAATTGAAGACAGAAACACAAATTAAGTCCAAAGCAATGGATTCTA 10059
QY 61 TTGGGAGTGAATTCGCCCTTAGAAGACACTGGCAATACCAAGGACATTTTGGTTCTCA 120
Db 10060 TTGGGAGTGAATTCGCCCTTAGAAGACACTGGCAATACCAAGGACATTTTGGTTCTCA 10119
QY 121 CAATATATGGAGGGGCAATTAAGTGGCAATTAAGTGGATAGATGCCAAGTGTCTTCAAC 180
Db 10120 CAATATATGGAGGGGCAATTAAGTGGCAATTAAGTGGATAGATGCCAAGTGTCTTCAAC 10179
QY 181 ATGCTATGATGCACAGGCGGCTCCAAACAAACCAATTAAGTGGATAGATGCCAAGTGTCTTCAAC 240
Db 10180 ATGCTATGATGCACAGGCGGCTCCAAACAAACCAATTAAGTGGATAGATGCCAAGTGTCTTCAAC 10239
QY 241 AGTCCCAGATCCAGGACCCCTCATCCAGGGCTGAGAACCGTATTTTCCAGAGGGAG 300
Db 10240 AGTCCCAGATCCAGGACCCCTCATCCAGGGCTGAGAACCGTATTTTCCAGAGGGAG 10299
QY 301 GTATAGGATGGTGTGGAGAAATGGGAAGAGGTGTGTCTCAGTAAAGAAATAA 360
Db 10300 GTATAGGATGGTGTGGAGAAATGGGAAGAGGTGTGTCTCAGTAAAGAAATAA 10359
QY 361 GGCTTGACAGGCTGGAGGGGAGAGTGGAGAGAAAGGGAGGGAGAGATACACGATGA 420
Db 10360 GGCTTGACAGGCTGGAGGGGAGAGTGGAGAGAAAGGGAGGGAGAGATACACGATGA 10419
QY 421 GGAGACAGGCTGGAAACAGAAAGTAGAGACAGAGATTTCGAGATGTGGAGAGGAAGGTCA 480
Db 10420 GGAGACAGGCTGGAAACAGAAAGTAGAGACAGAGATTTCGAGATGTGGAGAGGAAGGTCA 10479
QY 481 CAGACCCCGGAAATGATGTGGCAACACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 540
Db 10480 CAGACCCCGGAAATGATGTGGCAACACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 10539

QY 541 AGTGACAAATGGGGTCTAAAGGTTGAACTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 600
Db |||||
10540 AGTGACAAATGGGGTCTAAAGGTTGAACTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 10599
QY 601 ATCCCAACACTTTGGAGGCTGAGTGGCGCAATCACTTGGAGGCCAGGAGTTCGAGACCAG 660
Db |||||
10600 ATCCCAACACTTTGGAGGCTGAGTGGCGCAATCACTTGGAGGCCAGGAGTTCGAGACCAG 10659
QY 661 CTGGGCAACATGGTGAACCCCGCTCTCTCAAAAAAATACAAAAAATAGCCGGGTGT 720
Db |||||
10660 CTGGGCAACATGGTGAACCCCGCTCTCTCAAAAAAATACAAAAAATAGCCGGGTGT 10719
QY 721 GGTGATGGAACCTGTAGTACAGCTACTTGGAGGCTGAGGAGGAGGAATGCTTGAAC 780
Db |||||
10720 GGTGATGGAACCTGTAGTACAGCTACTTGGAGGCTGAGGAGGAGGAATGCTTGAAC 10779
QY 781 CCGGAGATGAGGCTGCAGTGAGCTGAGTCAAGCTGAGCCACCTGCGCTCCAACTGGGCAACA 840
Db |||||
10780 CCGGAGATGAGGCTGCAGTGAGCTGAGTCAAGCTGAGCCACCTGCGCTCCAACTGGGCAACA 10839
QY 841 GAGTAAGACTCCATCTCAAAAAAAGCTGATTTGGAGTGAATATTAATAACATT 900
Db |||||
10840 GAGTAAGACTCCATCTCAAAAAAAGCTGATTTGGAGTGAATATTAATAACATT 10899
QY 901 CTCCTCTCTCTCTTTTGGCTGTCTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT 960
Db |||||
10900 CTCCTCTCTCTCTTTTGGCTGTCTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT 10959
QY 961 CTGTACTTTCCATCTCTGTGTCTGTTCCTCACTGCTCTTCCATCTATGAGGATCTCTG 1020
Db |||||
10960 CTGTACTTTCCATCTCTGTGTCTGTTCCTCACTGCTCTTCCATCTATGAGGATCTCTG 11019
QY 1021 GGTCTCTCATGTCTCTTCTGCCACATCTTGGCAATCTCTGCTCTCTCATGCCCCCTT 1080
Db |||||
11020 GGTCTCTCATGTCTCTTCTGCCACATCTTGGCAATCTCTGCTCTCTCATGCCCCCTT 11079
QY 1081 TCTCTCTCGAGGATGATTTGGGGGCGCTGTGTCTGCAATGGCTCCCTCGAGGACTC 1140
Db |||||
11080 TCTCTCTCGAGGATGATTTGGGGGCGCTGTGTCTGCAATGGCTCCCTCGAGGACTC 11139
QY 1141 GTGTCTGGGAGATTTACCTTTGTCCCGGCCCAAGACGCGGTGTCTACAGAACCTC 1200
Db |||||
11140 GTGTCTGGGAGATTTACCTTTGTCCCGGCCCAAGACGCGGTGTCTACAGAACCTC 11199
QY 1201 TGCAAGTTTCAACAGTGGATCCAGAAACCATCCAGGCCAACTCCTGAGTCATCCCAAGGA 1260
Db |||||
11200 TGCAAGTTTCAACAGTGGATCCAGAAACCATCCAGGCCAACTCCTGAGTCATCCCAAGGA 11259
QY 1261 CTCAGCACACCGGCATCCCCCTGCTGCGAGGACAGCCCTGACACTCCTTTGAGACCTT 1320
Db |||||
11260 CTCAGCACACCGGCATCCCCCTGCTGCGAGGACAGCCCTGACACTCCTTTGAGACCTT 11319
QY 1321 CATTCCTTCCAGAGATGTTGAGATGTTATCTCTCCAGCCCTGACCCCATGCTCTCT 1380
Db |||||
11320 CATTCCTTCCAGAGATGTTGAGATGTTATCTCTCCAGCCCTGACCCCATGCTCTCTCT 11379
QY 1381 GGACTCAGGCTCTGCTTCCCCCATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGG 1440
Db |||||
11380 GGACTCAGGCTCTGCTTCCCCCATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGG 11439
QY 1441 AACAAATTTCCAAAACCTGTCCAGGGCGGGGTGGGTCTCAATCTCCCTGGGGCACTTTCA 1500
Db |||||
11440 AACAAATTTCCAAAACCTGTCCAGGGCGGGGTGGGTCTCAATCTCCCTGGGGCACTTTCA 11499
QY 1501 TCCTCAAGCTCAGGGCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAAT 1560
Db |||||
11500 TCCTCAAGCTCAGGGCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAAT 11559
QY 1561 AAACCTGAGAAG 1571
Db |||||
11560 AAACCTGAGAAG 11570

RESULT 4
AF135028
LOCUS 11570 bp DNA linear PRI 27-JUN-2000
DEFINITION Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
ACCESSION AF135028
VERSION AF135028.1 GI:4589282
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 11570)
AUTHORS Yousef, G.M. and Diamandis, E.P.
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,
mapping, tissue expression, and hormonal regulation
J. Biol. Chem. 274 (53), 37511-37516 (1999)
JOURNAL 20076408
MEDLINE 1060802
PUBMED
REFERENCE 2 (bases 1 to 11570)
AUTHORS Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Obiezu, C.V.
TITLE The new human kallikrein gene family: implications in
carcinogenesis
Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
JOURNAL 21121728
MEDLINE 10675891
PUBMED
REFERENCE 3 (bases 1 to 11570)
AUTHORS Yousef, G.M., Luo, L.Y. and Diamandis, E.P.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
FEATURES
source
1. .11570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101..2131,2210..2293,4762..5023,5763..6019,
6105..6238,11092..11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221..2293,4762..5023,5763..6019,6105..6238,
11092..11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"
/protein_id="AAD26429.1"
/db_xref="GI:4589283"
/translation="MAIARPPWVLCALITALLLVTEHVLANNVSCDHPSTVPS
GNSQDILGAGGEDARSDSSRIINGSDDMHTQPAQALLRPNDLYCGAVLVHPW
LITAAHCRKKVFRVLGHYSLSFVSGQMGQVKSIPHPYSHPGHSDMLIKLN
RRIRPKDVRPINVSSHCPKSLVSGMTKSLVQHPFKVLQCLINISLVLSOKCE
DAYPROIDTDFMPCAGDKAGRDSQCGDGGPVVCGNSLQGLVSGNDYPCARPNRPGVVT
NLCKFTWITQETIQANS"
mRNA
CDS
ORIGIN
Query Match 100.0%; Score 1571; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGAGGATGAGGAATTGAAGACAGAAACACAAATAGTCCAGGAATGGATTCTTA 60
Db 10000 AGGAGGATGAGGAATTGAAGACAGAAACACAAATAGTCCAGGAATGGATTCTTA 10059
QY 61 TTGGAGTGAATTCGCCCTTAGAAGACACCTGGCAATACAGGAGACATTTTGGTGTCA 120
Db 10060 TTGGAGTGAATTCGCCCTTAGAAGACACCTGGCAATACAGGAGACATTTTGGTGTCA 10119
QY 121 CAACATATATGAGGGGCAATTAAGCACTAATAGGATAGATGCCAAGTGTCTGTCAAC 180
Db 10120 CAACATATGAGGGGCAATTAAGCACTAATAGGATAGATGCCAAGTGTCTGTCAAC 10179
QY 181 ATGCTATGATGACACAGCGCGCCCTCCACAAACCAATATTCAGATTCAGATGCCAC 240

10180	Db	 ATGCTATGATGCACAGCGAGCGCTCCACAAACAAACCAATTATCCAGCTTCAGATGCCAC	10239
241	QY	 AGTGCCCGAGATCAGAGAAACCTCATCCAGGGCTCAGAACCGTATTTTTCAGAAAGGAG	300
10240	Db	 AGTGCCCGAGATCGAGGAAACCTCATCCAGGGCTCGAAACCGTATTTTTCAGAAAGGAG	10299
301	QY	 GTATAAGGATGGGTTCGTGGAGAAATGGGGAGGAAAGGTGTGTCCAGTAAGAGAAATAA	360
10300	Db	 GTATAAGGATGGGTTCGTGGAGAAATGGGGAGGAAAGGTGTGTCCAGTAAGAGAAATAA	10359
361	QY	 GGCTCGCACAGCTGGAGGGGAGATGTAGAGAAAGGGAGCGCGAGAGATACACGATGA	420
10360	Db	 GGCTCGCACAGCTGGAGGGGAGATGTAGAGAAAGGGAGCGCGAGAGATACACGATGA	10419
421	QY	 GGGAGACAGCTCGAAACAGAAAGTAGACAGCAAGATTCCAGATGTGGAGAGAGAGGGTCA	480
10420	Db	 GGGAGACAGCTCGAAACAGAAAGTAGACAGCAAGATTCCAGATGTGGAGAGAGAGGGTCA	10479
481	QY	 CAGACCCCCGAAATGTGTGGACAAACAGGAATCTGGAAGAGAAAGATGGAGTGGAG	540
10480	Db	 CAGACCCCCGAAATGTGTGGACAAACAGGAATCTGGAAGAGAAAGATGGAGTGGAG	10539
541	QY	 AGTGACAAATGGGTCCTAAAGTTGAACTTTGAGAGCCAGGCATGTGTGCTCACGCCTGTA	600
10540	Db	 AGTGACAAATGGGTCCTAAAGTTGAACTTTGAGAGCCAGGCATGTGTGCTCACGCCTGTA	10599
601	QY	 ATCCCAACACTTTGGAGGCTGAGGTGGGCGAATCACTTGAGCCAGGAGTTTCGAGACCAG	660
10600	Db	 ATCCCAACACTTTGGAGGCTGAGGTGGGCGAATCACTTGAGCCAGGAGTTTCGAGACCAG	10659
661	QY	 CTTGCCCAACATGGTGAACCCCGTCTCTACAAAAAAATACAAAAATTAGCCGGTGT	720
10660	Db	 CTTGCCCAACATGGTGAACCCCGTCTCTACAAAAAAATACAAAAATTAGCCGGTGT	10719
721	QY	 GGTGATGGACACTGTAGTCAAGCTACTTTGGAGGCTTGAGCAGGAGAAATTTGCTTGAC	780
10720	Db	 GGTGATGGACACTGTAGTCAAGCTACTTTGGAGGCTTGAGCAGGAGAAATTTGCTTGAC	10779
781	QY	 CCGGAGATGGAGGCTGCAGTGAGCTCAGGCGCACTGGGCTCCAACTGGGCAACA	840
10780	Db	 CCGGAGATGGAGGCTGCAGTGAGCTCAGGCGCACTGGGCTCCAACTGGGCAACA	10839
841	QY	 GAGTAAGACTCCATCTCAAAAAAAAAGCTGGATTTGGAGTGAATAATTAATACATT	900
10840	Db	 GAGTAAGACTCCATCTCAAAAAAAAAGCTGGATTTGGAGTGAATAATTAATACATT	10899
901	QY	 CTCCCTCTCTCCTTTTGGCTGTGTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT	960
10900	Db	 CTCCCTCTCTCCTTTTGGCTGTGTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT	10959
961	QY	 CTGTACTTTTCCATCTCTGTGTCTGTTCCTTCCCATCTGTCTCCATCTATGGCATCTCTG	1020
10960	Db	 CTGTACTTTTCCATCTCTGTGTCTGTTCCTTCCCATCTGTCTCCATCTATGGCATCTCTG	11019
1021	QY	 GGTCTCTCATGTCTCCTTTCTGCCCACTTTGGCCACATCTCTGCTCTCTCATGCCCCCTT	1080
11020	Db	 GGTCTCTCATGTCTCCTTTCTGCCCACTTTGGCCACATCTCTGCTCTCTCATGCCCCCTT	11079
1081	QY	 TCTCTCTGTGAGGTCATTTCTGGGGGCTCTGTGGTCTGCAATGGCTTCCTGTGAGGATC	1140
11080	Db	 TCTCTCTGTGAGGTCATTTCTGGGGGCTCTGTGGTCTGCAATGGCTTCCTGTGAGGATC	11139
1141	QY	 GTGTCCTGGGAGATTACCTTTGTGCCCGGCCCAACAGACCCGGGTGTCTACACGAACCTC	1200
11140	Db	 GTGTCCTGGGAGATTACCTTTGTGCCCGGCCCAACAGACCCGGGTGTCTACACGAACCTC	11199
1201	QY	 TGCAAGTTACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGA	1260
11200	Db	 TGCAAGTTACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGA	11259
1261	QY	 CTCAGCACACCGGCATCCCACTGTGTGAGGAGACGCCCTGACACTCTCTTTTCAGACCT	1320

Db	11260	CTCAGCACACCGGCATCCCACTGCTGCAGGGA	CAGCCCTGACACTCCTTTTCAGACCCCT	11319
Qy	1321	CATTCTCTCCCGAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTCGACCCCATGTCTCCT	1380	
Db	11320	CATTCTCTCCCGAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTCGACCCCATGTCTCCT	11379	
Qy	1381	GGACTCAGGCTGCTTTCCCCACATCTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGG	1440	
Db	11380	GGACTCAGGCTGCTTTCCCCACATCTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGG	11439	
Qy	1441	AACAATTTCCAAACATGTTCCAGGGGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCA	1500	
Db	11440	AACAATTTCCAAACATGTTCCAGGGGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCA	11499	
Qy	1501	TCCTCAAGCTCAGGGCCCATCCCTTCTCTGCGAGCTCTGACCCAAATTTAGTCCCGAAAT	1560	
Db	11500	TCCTCAAGCTCAGGGCCCATCCCTTCTCTGCGAGCTCTGACCCAAATTTAGTCCCGAAAT	11559	
Qy	1561	AAACTGAGAAG	1571	
Db	11560	AAACTGAGAAG	11570	
RESULT 5				
AC011483/c				
LOCUS				
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.				
ACCESSION AC011483				
VERSION AC011483.7				
KEYWORDS HTG.				
SOURCE Homo sapiens (human)				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
misc_feature				
:				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 1571; Conservative				
100.0%; Score 1571; DB 9; Length 107487;				
100.0%; Pred. No. 0;				
0; Mismatches				
0; Indels				
0; Gaps				
0;				

Qy	1	AGGGAGGTATGGGGAAATTGAAGACAGGAAACACAAATTTAGTTCAAAGCGAATGGATTTCTTA	60
Db	50220	AGGGAGGTATGGGGAAATTGAAGACAGGAAACACAAATTTAGTTCAAAGCGAATGGATTTCTTA	50161
Qy	61	TTGGGAGTGATTTCTGCCCTTGAAGAACACTGGCAATACCAGAGACATTTTTTGTTGTGTCA	120
Db	50160	TTGGGAGTGATTTCTGCCCTTGAAGAACACTGGCAATACCAGAGACATTTTTTGTTGTGTCA	50101
Qy	121	CAACTATATGAGGGGCAATTACTTGGCAACTAATTGGATAGATGCCAAGTGTGCTTTC AAC	180
Db	50100	CAACTATATGAGGGGCAATTACTTGGCAACTAATTGGATAGATGCCAAGTGTGCTTTC AAC	50041
Qy	181	ATGCTATATGACACAGGCAAGGCTTCCACAAACCAATTTATCCAGCTTCAGATGCCAC	240
Db	50040	ATGCTATATGACACAGGCAAGGCTTCCACAAACCAATTTATCCAGCTTCAGATGCCAC	49981
Qy	241	AGTGCCCAGATTCAGGAACCCCTCATCCAGGGCTGAGAACCCGTATTTTTTGCAGAGGAG	300
Db	49980	AGTGCCCAGATTCAGGAACCCCTCATCCAGGGCTGAGAACCCGTATTTTTTGCAGAGGAG	49921
Qy	301	GTATAAGGATGGGTTGGTGAGAAATCGGAAGAAAGGTGTGTCCAGTAAGAGAAATAA	360
Db	49920	GTATAAGGATGGGTTGGTGAGAAATCGGAAGAAAGGTGTGTCCAGTAAGAGAAATAA	49861
Qy	361	GGCTTGCAAGGCTGGAGGGGAGGTGAGAGAAAGGAGCGGAGAGATACACATGA	420
Db	49860	GGCTTGCAAGGCTGGAGGGGAGGTGAGAGAAAGGAGCGGAGAGATACACATGA	49801
Qy	421	GGGAGACAGGCTCGAAACAGAAAGCTAGAGACGAAGATTCCAGATGTGGAGAGGAAGGTC	480
Db	49800	GGGAGACAGGCTCGAAACAGAAAGCTAGAGACGAAGATTCCAGATGTGGAGAGGAAGGTC	49741
Qy	481	CAGACCCCCCGGAAATGATGTGTGGAACAACAGAAATCTCGAAAGAGAAATGAGTGGAG	540
Db	49740	CAGACCCCCCGGAAATGATGTGTGGAACAACAGAAATCTCGAAAGAGAAATGAGTGGAG	49681
Qy	541	AGTGACAAATGGGGTCTTAAAGTTTGAACCTTGGAGGCCAGGCATGTGGCTCAGCCTGTA	600
Db	49680	AGTGACAAATGGGGTCTTAAAGTTTGAACCTTGGAGGCCAGGCATGTGGCTCAGCCTGTA	49621
Qy	601	ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTCAGGCACAGGAGTTTCGAGACCAG	660
Db	49620	ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTCAGGCACAGGAGTTTCGAGACCAG	49561
Qy	661	CTTGGCCAAACATGGTGAACCCCGTCTCTCAAAAAAATACAAAAAATTAGCCGGGTGT	720
Db	49560	CTTGGCCAAACATGGTGAACCCCGTCTCTCAAAAAAATACAAAAAATTAGCCGGGTGT	49501
Qy	721	GGTGATGGACACTGTAGTACACAGCTACTTGGGAGGCTGAGCAGCGAGAAATGCTTGAAC	780
Db	49500	GGTGATGGACACTGTAGTACACAGCTACTTGGGAGGCTGAGCAGCGAGAAATGCTTGAAC	49441
Qy	781	CCGGGAGATGGAGGCTGCAGTGAGCTGAGGTGAGGCCACTGGCTCCAACTCGGGCAACA	840
Db	49440	CCGGGAGATGGAGGCTGCAGTGAGCTGAGGTGAGGCCACTGGCTCCAACTCGGGCAACA	49381
Qy	841	GAGTAAGACTCCATCTCTCAAAAAAAGAGCTGGATTTGGAGTGAATAATTAATAACATT	900
Db	49380	GAGTAAGACTCCATCTCTCAAAAAAAGAGCTGGATTTGGAGTGAATAATTAATAACATT	49321
Qy	901	CTCCCTCTCTCCTTTTGGCTGTGTCTCCATCTCTGTCTTTTTCTGCAATTTCTTCATCT	960
Db	49320	CTCCCTCTCTCCTTTTGGCTGTGTCTCCATCTCTGTCTTTTTCTGCAATTTCTTCATCT	49261
Qy	961	CTGTACTTTTCCATCTCTGTGTGTCTGTTCCTCACTCTCTTCTCCATCTATGGGCATCTCTG	1020
Db	49260	CTGTACTTTTCCATCTCTGTGTGTCTGTTCCTCACTCTCTTCTCCATCTATGGGCATCTCTG	49201
Qy	1021	GGTCTCTCATGTCTCTTCTGCCCCACTTTTGCCACATCTCTGTGCTCTCTCATGCCCCCCTT	1080
Db	49200	GGTCTCTCATGTCTCTTCTGCCCCACTTTTGCCACATCTCTGTGCTCTCTCATGCCCCCCTT	49141

Qy	1081	TCTCTCTCGAGGGTGATTTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTCGAGGACTC	1140
Db	49140	TCTCTCTCGAGGGTGATTTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTCGAGGACTC	49081
Qy	1141	GTGTCTCGGGGAGATTACCTTTGTGCCCGGCCCAACAGACACGGGTGTCTACACGAACCTC	1200
Db	49080	GTGTCTCGGGGAGATTACCTTTGTGCCCGGCCCAACAGACACGGGTGTCTACACGAACCTC	49021
Qy	1201	TGCAAGTTCCAAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTAGTCAATCCAGGA	1260
Db	49020	TGCAAGTTCCAAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTAGTCAATCCAGGA	48961
Qy	1261	CTCAGCACACCGGCATCCCACTCTGTGTCAGGAGACGCGCTGACACTCTCTTTTCAGACCT	1320
Db	48960	CTCAGCACACCGGCATCCCACTCTGTGTCAGGAGACGCGCTGACACTCTCTTTTCAGACCT	48901
Qy	1321	CATTCTCTCCACAGATGTTGAGAATGTTTCATCTCTCCAGCCCTCGACCCCATGTCTCT	1380
Db	48900	CATTCTCTCCACAGATGTTGAGAATGTTTCATCTCTCCAGCCCTCGACCCCATGTCTCT	48841
Qy	1381	GGACTCAGGCTCTGCTTCTCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGG	1440
Db	48840	GGACTCAGGCTCTGCTTCTCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGG	48781
Qy	1441	AACAAATTTCCAAACATGTCACGGCGGGGTTTGGCTCAATCTCCCTGGGGCACTTTCA	1500
Db	48780	AACAAATTTCCAAACATGTCACGGCGGGGTTTGGCTCAATCTCCCTGGGGCACTTTCA	48721
Qy	1501	TCTCAAGCTCAGGGCCCATTCCTTCTTCGAGCTCTGACCCAAATTTAGTCCCAAAAT	1560
Db	48720	TCTCAAGCTCAGGGCCCATTCCTTCTTCGAGCTCTGACCCAAATTTAGTCCCAAAAT	48661
Qy	1561	AAACTGAGAAG	1571
Db	48660	AAACTGAGAAG	48650
RESULT 6			
AF243527/c			
LOCUS	AF243527	230000 bp	DNA linear PRI 21-NOV-2000
DEFINITION	Homo sapiens serine protease gene cluster, complete sequence.		
ACCESSION	AF243527		
VERSION	AF243527.1	GI:11244757	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 230000)		
AUTHORS	Moss, P., Paepel, B., and Wang, K.		
TITLE	Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region		
JOURNAL	Gene 257 (1), 119-130 (2000)		
MEDLINE	20510030		
PUBMED	11054574		
AUTHORS	Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J., Moss, P., Paepel, B. and Wang, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE, Bothell, WA 98021, USA		
FEATURES	Location/Qualifiers		
source	1. .230000		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map="19q13"		
gene	complement(<6963. .>11517)		
mRNA	complement(join(<6963. .7118,7668. .7804,7923. .8212, 9482. .9641,11472. .>11517))		


```
/gene="KLK1"
/product="renal kallikrein"
/complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/db_xref="GI:11244761"
/notes="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG33353.1"
/db_xref="GI:11244758"
/translation="MMFLVLCALSLSGTGGAAPPQISRIIVGVGWECEKHSQPWQVLAALYH
FSTFQCGIIVHQWLTAAHCISDNVQLWGRHNLFDDENTAQFVHVSFPHPGFN
MSLENNRDLQVDSHDLMLRLSEPAKTIIDAVKVELPTPEPEVSGTCLASGWSG
IEPENSFPDLOCDVLDKILPNDECKKAHVQKVTDEWLGVHLEGGKDTCTGDSGGPL
MCGVLOGVTSGWVPCGTENKPSVAVRVLSYVKWIEDTIAENS"
complement(join(<13352..13704,14377..14530,14678..14917,
15416..15569,19204..19246))
/product="ACO protease"
/complement(join(13552..13704,14377..14530,14678..14917,
15416..15569,19204..19246))
/notes="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MWLLLTSLFLLASTAODGDKLLEGDECAPHSQPWQVVALYERGR
FNGCASLISPHWLSAAHCOSRMVRVLGRHNLKRRDGPQLRTTSRVIPHPRYEARS
HRNDIMLLRLVQPARLPQVRPAVLPTRCFHPGEACVSGWGLVPLSSVSPVDPDTLHC
ANISIIDSTSCDYSYGRLTNTMVCAGAEAGSCEBGDSGGPLVCGGILQGVISWGD
VPCDNTTKPGVYTKVCHYLEWIRETKRN"
>42595..>47769
/gene="KLK3"
join(<42595..42640,43880..44039,45669..45955,46099..46235,
47614..>47769)
/gene="KLK3"
/product="prostate specific antigen"
join(<42595..42640,43880..44039,45669..45955,46099..46235,
47614..>47769)
/notes="serine protease"
/codon_start=1
/product="prostate specific antigen"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translation="MMVVFVFLTSVTWIGAAPLILSRIVGVGWECEKHSQPWQVLAAS
RGRAVCGVIVHPQWLTAAHCIRNKSIVILLGRHSLFHPEDTQGVQVSHSFPHPLYD
MSLLKNRFLRPDGDSSHDLMLRLSEPAELTDAVKVMDLPTQBPALGTTTCYASGWGSI
EPFEFLTPKLQCVLDLHVISNDVCAOVHPQKVKEMLCAGRWTGKSTCGSDSGGPLV
CNGVLOGITSWSEPCALPERPSLYTKVHYRWIKDTIVANP"
<61139..>66229
/gene="KLK2"
join(<61139..61184,62391..62550,64142..64428,64542..64678,
66074..>66229)
/gene="KLK2"
/product="glandular kallikrein"
join(61139..61184,62391..62550,64142..64428,64542..64678,
66074..>66229)
/notes="serine protease"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MMWDLVLSIALSVGTGAVPLIQSRIIVGVGWECEKHSQPWQVAVVS
HGWAHCGVIVHPQWLTAAHCLKNSQVWLGRHNLPEPEDTQGVQVSHSFPHPLYN
MSLLKHOSLPDEDSHDLMLRLSEPAKTIIDVVKVLGLTQBPALGTTTCYASGWGSI
EPFEFLRPSLQCVLSHLLNDMCARAYSEKVTFLCAGLWTKGKDTCCGDSGGPLV
CNGVLOGITSWGEPCALPEKPAVYTKVHYRWIKDTIAANP"
complement(join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..>98405))
/product="serine protease"
complement(join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..>98405))
```

```
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPMGWFLGYLILVAGSLVSGSCSIINGEDCSPHSOPW
QAALWENELFCGVLVHPQWLSAAHCFQNSYTIGLHLHLEADQEPGSOYVEASLS
VRHPEYRPLLANDMLIKLDESVSSEDRTISIASQCFTAGNSCLVSGWGLLANGR
MPTVLOCVNVSUVEVSKLYDPLYPHSMFCAGGHDQKDCSNGSGGPLICNGYLO
GLVSPGAPCGQGVGVYTNLCKFTETWIKTQVQS"
complement(join(131301..>140303))
/notes="KLK5"
/notes="synonym: SCTE"
complement(join(<131301..131456,136310..136443,
136529..136785,137525..137690,140255..>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement(join(131301..131456,136310..136443,
136529..136785,137525..137690,140255..>140303))
/notes="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MMVLCALITALLGYTDARSDDSSRIINGSDCDMHTQPWQAL
LARPOLYCGAVLVHPQWLLTAACHCKKVRVRLGHVLSLPVYESGOQFQVKSIPH
PGYSHCHNDMLIKLNRIRPTKVDVINSSHCHPSAGTKCLVSGWGTTKSPQVHF
PKVLQCLNLSVLSQKCEDAYPQIDDTMFCAGDKAGRDCSQGDSGGVVCNGSLQGL
VWNGDYPCARPNNRPGVYTNLCKFTWIKTQVQS"
complement(join(<146834..146986,149628..149764,
151186..151433,155052..155208,155948..>155987))
/product="protease M"
complement(join(146834..146986,149628..149764,
151186..151433,155052..155208,155948..>155987))
/notes="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKKLMVLSLIAAAWAEQNLVHGPGCDKTSHPYQAALYTSQH
LLCGVLIHPLWLTAACHCKNQLQVFLGKHNLRQRESSQSSVVRVAVIHPDYDAAS
HDQDMLRLARPAKLSLTIPLRLSDCSANTSCILGWGKTGDFPDPTIQCAVI
HLVRECECHAVPGQITONMLCAGDEKYGKDCSQGDSGGPLVCGDHLGLVSGWGNIPC
GSKEKPGVTNVCRTYTNWIKTQVQS"
complement(<165420..>170283)
/notes="KLK7"
/notes="synonym: SCCE"
complement(join(<165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement(join(165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283))
/notes="KLK7"
/notes="serine protease"
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MARSLLLPLQILLLSLALETAGEAOGDKIIDGAPCARGSHPWQ
VALLSGNLHCGVLNERWLTAAHCKMNEYTVHLGSDTLGDRRAQRIKASKSRFHP
GYSTQTHVNDMLVKNLSQARLSNNVKKVRLPSRCPPTCTVSGWGTTSPTDVTFP
SDMLCVDPKLNLSIPDQCTKVYKDLLENLMCAGIPDRSKKNACNDSGGPLVCRGLQGL
VWNGTFCPCQNDPGVYTVCKFTWINDTKKXH"
complement(join(<183943..184098,185635..185768,
187865..188127,188293..188452,188967..>189036))
/product="neuropsin"
complement(join(183943..184098,185635..185768,
187865..188127,188293..188452,188967..>189036))
/notes="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
```


REFERENCE AUTHORS TITLE	JOURNAL REFERENCE AUTHORS
1. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	1. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
2. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	2. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
3. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	3. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
4. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	4. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
5. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	5. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
6. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	6. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
7. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	7. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
8. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	8. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
9. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	9. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001
10. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001	10. <i>Journal of the American Medical Association</i> , 1964, 191: 1000-1001

1 (bases 1 to 217346)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 19, clone RP11-795B6
Unpublished

2 (bases 1 to 217346)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukngalter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Chaoelano, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dechrell, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuck, A., McKernan, K., MCPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Olivat, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, S., Riley, K., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 217346)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

```

Insert size: 194000; agarose-fp
Insert size: 216246; sum-of-ontigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.

NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	1	8149:	contig of 8149 bp in length
*	8150	8249:	gap of 100 bp
*	8250	9592:	contig of 1343 bp in length
*	9593	9692:	gap of 100 bp
*	9693	10733:	contig of 1041 bp in length
*	10734	10833:	gap of 100 bp
*	10834	13519:	contig of 2686 bp in length
*	13520	13619:	gap of 100 bp
*	13620	17510:	contig of 3891 bp in length
*	17511	17610:	gap of 100 bp
*	17611	24602:	contig of 6992 bp in length
*	24603	24702:	gap of 100 bp
*	24703	35434:	contig of 10732 bp in length
*	35435	35534:	gap of 100 bp
*	35535	124474:	contig of 88940 bp in length
*	124475	124574:	gap of 100 bp
*	124575	134664:	contig of 10090 bp in length
*	134665	134764:	gap of 100 bp
*	134765	162343:	contig of 27579 bp in length
*	162344	162443:	gap of 100 bp
*	162444	208917:	contig of 46474 bp in length
*	208918	209017:	gap of 100 bp
*	209018	217346:	contig of 8329 bp in length.

FEATURES
source

```

source
1. .217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_lib="RPC1-11 Human Male BAC"
1. .8149
/note="assembly_fragment
clone end:SP6
vector side:left"
8250. .9592
/note="assembly_fragment"
9693. .10733
/note="assembly_fragment"
10834. .13519
/note="assembly_fragment"
13620. .17510
/note="assembly_fragment"
17611. .24602
/note="assembly_fragment"
24703. .35434
/note="assembly_fragment"
35535. .124474
/note="assembly_fragment"
124575. .134664
/note="assembly_fragment"
134765. .162343
/note="assembly_fragment"
162444. .208917
/note="assembly_fragment"
209018. .217346
/note="assembly_fragment
clone end:T7
vector side:right"

```

ORIGIN

Query Match		99.0%; Score 1555.2; DB 2; Length 217346;
Best Local Similarity		99.7%; Pred. No. 0;
Matches 1568; Conservative		0; Mismatches 3; Indels 1; Gaps 1;
QY	1	AGGAGGTATGGGAAATTGAAGACAGAGAAACACAAATTAAGTCCAAAGCAATGATTTCTA 60
DB	194423	AGGAGGTATGGGAAATTGAAGACAGAGAAACACAAATTAAGTCCAAAGCAATGATTTCTA 194364
QY	61	TTGGGAGTATTCGCCCTTAGAGACACTGGCNAATACAGGAGACATTTTGGTGTCA 120
DB	194363	TTGGGAGTATTCGCCCTTAGAGACACTGGCNAATACAGGAGACATTTTGGTGTCA 194304
QY	121	CAACTATATGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAAGTGTGTTCAAC 180
DB	194303	CAACTATATGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAAGTGTGTTCAAC 194244
QY	181	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATTTATCCAGTTCAGATGCCAC 240
DB	194243	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATTTATCCAGTTCAGATGCCAC 194184
QY	241	AGTCCCGAGATCGAGGAACCTCATCCAGGGCTGAAACGCTATTTTTCAGAAAGGAG 300
DB	194183	AGTCCCGAGATCGAGGAACCTCATCCAGGGCTGAAACGCTATTTTTCAGAAAGGAG 194124
QY	301	GTATAAGGATGGGTGGTGAGAAATGGGAAGGAGGTGTGTCCAGTAAGAGAAATAA 360
DB	194123	GTATAAGGATGGGTGGTGAGAAATGGGAAGGAGGTGTGTCCAGTAAGAGAAATAA 194064
QY	361	GGCTGACAGGCTGGAGGGAGAGTGAAGAGAAAGGAGGGCGGAGAGATACACGATGA 420
DB	194063	GGCTGACAGGCTGGAGGGAGAGTGAAGAGAAAGGAGGGCGGAGAGATACACGATGA 194004
QY	421	GGGAGACAGGCTGGAAACAGAAAGTAGAGACGAAGATTCGAGATGTGGAGAGAGGTCA 480
DB	194003	GGGAGACAGGCTGGAAACAGAAAGTAGAGACGAAGATTCGAGATGTGGAGAGAGGTCA 193944
QY	481	CAGACCCCGGAAATGATGTGGCAACACAGGAATCTGGAAGAGGAAGATGAGTGGAG 540
DB	193943	CAGACCCCGGAAATGATGTGGCAACACAGGAATCTGGAAGAGGAAGATGAGTGGAG 193884
QY	541	AGTGACAAATGGGTCTAAAGTTGAACTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 600
DB	193883	AGTGACAAATGGGTCTAAAGTTGAACTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 193824
QY	601	ATCCCAACACTTGGAGGCTGAGTGGCGAATCATTGAGGCCAGGAGTTCGAGACCAG 660
DB	193823	ATCCCAACACTTGGAGGCTGAGTGGCGAATCATTGAGGCCAGGAGTTCGAGACCAG 193764
QY	661	CCTGGCCAAACATGGTGAACCCCGTCTCTACAAAAAAATACAAAAAATTTAGCCGGTGT 720
DB	193763	CCTGGCCAAACATGGTGAACCCCGTCTCTACAAAAAATTTAGCCGGTGT 193704
QY	721	GGTGATGGACACCTGTAGTACAGCTACTTGGGAGGCTGAGGAGGAGAAATTCCTGAAC 780
DB	193703	GGTGATGGACACCTGTAGTACAGCTACTTGGGAGGCTGAGGAGGAGAAATTCCTGAAC 193644
QY	781	CCGGAGATGGAGGCTGCAGTGAGCTGAGTTCAGGCCACTGCGCTCCAACTGGGCAACA 840
DB	193643	CTGGGAGATGGAGGCTGCAGTGAGCTGAGTTCAGGCCACTGCGCTCCAACTGGGCAACA 193584
QY	841	GAGTAAGACTCCATCTC-AAAAAAAAAAAAAGCTGGATTTGGAGTGAATAATTAATAACAT 899
DB	193583	GAGTAAGACTCCATCTC-AAAAAAAAAAAAAGCTGGATTTGGAGTGAATAATTAATAACAT 193524
QY	900	TCTCCCTCTCTCTCTTTTGGCTGTGTCTGCATCTCTGTCTTTTTCGCAATTTCTCATC 959
DB	193523	TCTCCCTCTCTCTCTTTTGGCTGTGTCTGCATCTCTGTCTTTTTCGCAATTTCTCATC 193464
QY	960	TCTGTACTTTCCATCTCTGTGTGTCTGTTCCTCCATCTCTCTCTATATGGGCATCTCT 1019
DB	193463	TCTGTACTTTCCATCTCTGTGTGTCTGTTCCTCCATCTCTCTCTATATGGGCATCTCT 193404

RESULT 8
AC130782/c

LOCUS	AC130782	200792 bp	DNA	linear	HTG 22-NOV-2002
DEFINITION	Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12 ordered pieces.				
ACCESSION	AC130782				
VERSION	AC130782.2	GI:25167101			
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.				
SOURCE	Pan troglodytes (chimpanzee)				
ORGANISM	Pan troglodytes				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.				
REFERENCE	1 (bases 1 to 200792)				
AUTHORS	Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lalic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., McDowell, J., Margulies, E.H., Masiello, C., Maskeri, B., Portnoy, M.E., Prasad, A., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddi-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Standrip, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.				
TITLE	NISC Comparative Sequencing Initiative				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 200792)				
AUTHORS	Green, E.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA				
REFERENCE	3 (bases 1 to 200792)				
AUTHORS	Green, E.D.				

TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT On Nov 22, 2002 this sequence version replaced gi:22218452.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoelnigri.nih.gov
----- Project Information
Center project name: dhz
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 28306: contig of 28306 bp in length

* 28307 28406: gap of unknown length

* 28407 37858: contig of 9450 bp in length

* 37857 37956: gap of unknown length

* 37957 73522: contig of 35566 bp in length

* 73523 73622: gap of unknown length

* 73623 83567: contig of 9945 bp in length

* 83568 83667: gap of unknown length

* 83668 88817: contig of 5150 bp in length

* 88817 88917: gap of unknown length

* 88917 125611: contig of 36694 bp in length

* 125611 125711: gap of unknown length

* 125712 159879: contig of 34168 bp in length

* 159880 159979: gap of unknown length

* 159980 174698: contig of 14719 bp in length

* 174699 174798: gap of unknown length

* 174799 186382: contig of 11584 bp in length

* 186383 186482: gap of unknown length

* 186483 193344: contig of 6862 bp in length

* 193345 193445: gap of unknown length

* 193445 199363: contig of 5919 bp in length

* 199364 199463: gap of unknown length

* 199464 200792: contig of 1329 bp in length.

FEATURES
source Location/Qualifiers
1..200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"

/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"
1..28306
/notes="assembly_fragment"
clone_end:SP6
vector_side:left
28407..37856
/notes="assembly_fragment"
37957..73522
/notes="assembly_fragment"
73623..83567
/notes="assembly_fragment"
83668..88817
/notes="assembly_fragment"
88918..125611
/notes="assembly_fragment"
125712..159879
/notes="assembly_fragment"
159980..174698
/notes="assembly_fragment"
174799..186382
/notes="assembly_fragment"
186483..193344
/notes="assembly_fragment"
193445..199363
/notes="assembly_fragment"
199464..200792
/notes="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN

Query Match 94.3%; Score 1482; DB 2; Length 200792;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1542; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

QY 1 AGGAGGTATGGGAAATTGAAGACAGGAACACAAATTAGTCCNAGGAATGGATTCTTA 60
Db 69290 AGGAGGTATGGGAAATTGAAGACAGGAACACAAATTAGTCCNAGGAATGGATTCTTA 69231

QY 61 TTGGGAGTGTATCTGCCCCCTAGAGACACTGGCAATACACAGGAGACATTTTGGTGTCA 120
Db 69230 TTGGGAGTGTATCTGCCCCCTAGAGACACTGGCAATACACAGGAGACATTTTGGTGTCA 69171

QY 121 CAATATATGGAGGGGCATTACTGGCAACTAATAGATAGATGCAAGTGTCTGCTTCAAC 180
Db 69170 CAATATATGGAGGGGCATTACTGGCAACTAATAGATAGAGGCCAAGGTGCTTCAAC 69111

QY 181 ATGCTATGATGCACGGCAGGCTTCCAAACAAACCAATTCAGTTCAGATGCCAC 240
Db 69110 GTGCTATGATGCACGGCAGGCTTCCAAACAAACCAATTCAGTTCAGATGCCAC 69051

QY 241 AGTGGCCAGATCCAGGNAACCTCATCCAGGGCTGAGAACCGTATTTTCCAGNAGGAG 300
Db 69050 AGTGGCCAGATCCAGGNAACCTCATCCAGGGCTGAGAACCGTATTTTCCAGNAGGAG 68991

QY 301 GTATAGAGATCGGTTGGTGAGAAATGGGAAGCAAGTGTGTGCCAGTAAGAGAAATAA 360
Db 68990 GTATAGAGATCGGTTGGTGAGAAATGGGAAGCAAGTGTGTGCCAGTAAGAGAAATAA 68931

QY 361 GGCCTGCACAGGTGGAGGGGAGTGAGAGAGAAAGGGAGGGGAGAGATACACGATGA 420
Db 68930 GGCCTGCACAGGTGGAGGGGAGTGAGAGAGAAAGGGAGGGGAGAGATACACGATGA 68871

QY 421 GGGAGACAGCTGGACAGAACTAGACAGAGATTCGAGATGTGAGAGGAGGGTCA 480
Db 68870 GGGAGACAGCTGGACAGAACTAGACAGAGATTCGAGATGTGAGAGGAGGGTCA 68811

QY 481 CAGACCCCCCGGAAATGATGTGTGGACACACAGGAATCTGGAAGAGGAGAGTGGAGTGGAG 540
Db 68810 CAG-GCCCCCGGAAATGATGTGT-GACACAGGAATCTGGAAGAGGAGAGTGGAGTGGAG 68753

QY 541 AGTGACAAATGGGTCTAAAGTTGAACCTTGGAGCCAGGATGGTGGCTCAGCCCTGTA 600
Db |||||
68752 AGTGACAAATGGGTCTGGAAGTTGAACCTTGGAGCCAGGATGGTGGCTCAGCCCTGTA 68693
QY 601 ATCCCAACATTTGGAGGCTGAGTGGGCGAATCATTGAGGCCAGGAGTTGAGACCAG 660
Db |||||
68692 ATCCCAATACCTTTGGAGGCTGAGTGGGCGAATCATTGAGGCCAGGAGTTGAGACCAG 68633
QY 661 CTTGGCCAAATGATGTAACCCCGTCTCTAC-AAAAAAAATACAAAAAATAGCCGGTG 719
Db |||||
68632 CTTGGCCAAATGATGTAACCCCGTCTCTACAAAAAATACAAAAAATAGCCGGTG 68573
QY 720 TGTGATGACACCTGTAGTACAGCTACTTGGAGGCTGAGGAGGAGAGATGCTTGAA 779
Db |||||
68572 TGTGATGACACCTGTAGTACAGCTACTTGGAGGCTGAGGAGGAGAGATGCTTGAA 68513
QY 780 CCGGGAGATGGAGGCTGAGTGGGCGAATCATTGAGGCCAGGAGTTGAGACCAG 839
Db |||||
68512 CCGGGAGATGGAGGCTGAGTGGGCGAATCATTGAGGCCAGGAGTTGAGACCAG 68453
QY 840 AGAGTAAGACTCCATCTCAAAAAAAGCTGGATTTGGAGTGAATATTAATACAT 899
Db |||||
68452 AGAGTAAGACTCCATCTC-AAAAAAGCTGGATTTGGAGTGAATATTAATACAT 68395
QY 900 TCTCCCTCTCTCTCTTTTGGCTGTCTCATCTCTGTCTCTTTTCTGCAATTTCTTCATC 959
Db |||||
68394 TCTCCCTCTCTCTCTTTTGGCTGTCTCTCTCTCTCTCTCTTTTCTGCAATTTCTTCATC 68335
QY 960 TCTGTACTTTCCATCTCTGTGTCTGTTCCTCATCTGCTTCTCATCTATGGGATCTCT 1019
Db |||||
68334 TCTGTACTTTCCATCTCTGTGTCTGTTCCTCATCTGCTTCTCATCTATGGGATCTCT 68275
QY 1020 GGGTCTCTCATCTCTCTCTGCCCACCTTTGGCCACATCTCTGCTCTCTCATGCCCT 1079
Db |||||
68274 GGGTCTCTCATCTCTCTCTGCCCACCTTTGGCCACATCTCTGCTCTCTCATGCCCT 68215
QY 1080 TTCTCTCTCTGAGGTTGATTTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCT 1139
Db |||||
68214 TTCTCTCTGAGGTTGATTTGGGGGCTGTGGTCTGCAACGGCTCCCTGAGGAGCT 68155
QY 1140 CGTGTCTCTGAGGATTTACCTTTGGCCGCGCCCAACAGACCAGGTTCTTACAGAACCT 1199
Db |||||
68154 CGTGTCTCTGAGGATTTACCTTTGGCCGCGCCCAACAGACCAGGTTCTTACAGAACCT 68095
QY 1200 CTGCAAGTTCCACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCTATCCAGG 1259
Db |||||
68094 CTGCAAGTTCCACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCTATCCAGG 68035
QY 1260 ACTCAGCACACCGGATCTCCACCTGCTGAGGAGCAGCCCTGACACTCTCTTTTACAGACC 1319
Db |||||
68034 ACTCAGCACACCGGATCTCCACCTGCTGAGGAGCAGCCCTGACACTCTCTTTTACAGACC 67975
QY 1320 TCATTCTTCCAGAGATGTTGAGATGTTCACTCTCTCAGCCCTGACCCCATGTCCTCC 1379
Db |||||
67974 TCATTCTTCCAGAGATGTTGAGATGTTCACTCTCTCAGCCCTGACCCCATGTCCTCC 67915
QY 1380 TGAGTCTGAGGTTCTTCCCCACATTTGGCTGACCGTGTCTCTCTAGTTGAACCTG 1439
Db |||||
67914 AGGACTCAGGTTCCGCTTCCCCACATTTGGCTGACCGTGTCTCTCTAGTTGAACCTG 67855
QY 1440 GAACAAATTTCCAAATCTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGAGGACATTTTC 1499
Db |||||
67854 GAACAAATTTCCAAATCTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGAGGACATTTTC 67795
QY 1500 ATCTCAAGCTCAGGCGCCATCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAA 1559
Db |||||
67794 ATCTCAAGCTCAGGCGCCATCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAA 67735
QY 1560 TAAACTGAGAG 1571
Db |||||
67734 TAAACTGAGAG 67723

RESULT 9
AX331407/c
LOCUS AX331407 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 1916 from Patent WO0194629.
ACCESSION AX331407
VERSION AX331407.1 GI:18122041
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 1916 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
source 1..586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 3.8e-131; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0;
QY 1090 CAGGGTGATTTGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGTCTGTCCTGG 1149
Db |||||
501 CAGGGTGATTTGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGTCTGTCCTGG 442
QY 1150 GGAGATTACCTTGTCCCGGCCCAACAGACCAGCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db |||||
441 GGAGATTACCTTGTCCCGGCCCAACAGACCAGCGGGTGTCTACAGAACCTCTGCAAGTTC 382
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269
Db |||||
381 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 322
QY 1270 CCGGCATCCCCACTCTCTGAGGAGCAGCCCTGACACTCTCTTTTTCAGACCTCTATCTTC 1329
Db |||||
321 CCGGCATCCCCACTCTCTGAGGAGCAGCCCTGACACTCTCTTTTTCAGACCTCTATCTTC 262
QY 1330 CCAGAGATGTTGAGATGTTTCACTCTCTCAGCCCTGACCCCATGTCCTCTGACTCAGG 1389
Db |||||
261 CCAGAGATGTTGAGATGTTTCACTCTCTCAGCCCTGACCCCATGTCCTCTGACTCAGG 202
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTTGGAAACAATTC 1449
Db |||||
201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTTGGAAACAATTC 142
QY 1450 CAAAACTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 1509
Db |||||
141 CAAAACTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 82
QY 1510 TCAGGCGCCATCTCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAACTGAGA 1569
Db |||||
81 TCAGGCGCCATCTCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAACTGAGA 22
QY 1570 AG 1571
Db || AG 20

RESULT 10
AX331824/c
LOCUS AX331824 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2333 from Patent WO0194629.
ACCESSION AX331824
VERSION AX331824.1 GI:18122458
KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horridan, S., Soppet, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 2333 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source
 1. .586
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 30.7%; Score 482; DB 6; Length 586;
 Best Local Similarity 100.0%; Pred. No. 3.8e-131; Mismatches 0; Indels 0; Gaps 0;
 Matches 482; Conservative 0;
 QY 1090 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCTGG 1149
 DB 501 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCTGG 442
 QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
 DB 441 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 382
 QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269
 DB 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 322
 QY 1270 CCGGCATCCCACTGTCTCCAGGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC 1329
 DB 321 CCGGCATCCCACTGTCTCCAGGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC 262
 QY 1330 CCAGAGATGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 1389
 DB 261 CCAGAGATGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 202
 QY 1390 GTCTGCTTCCCACTGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 1449
 DB 201 GTCTGCTTCCCACTGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 142
 QY 1450 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCTCGGGGACTTTCATCTCAAGC 1509
 DB 141 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCTCGGGGACTTTCATCTCAAGC 82
 QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
 DB 81 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 22
 QY 1570 AG 1571
 DB 21 AG 20
 RESULT 11
 G41906/c
 LOCUS
 DEFINITION SHGC-56840 Human Homo sapiens STS genomic, sequence tagged site.
 ACCESSION G41906
 VERSION G41906.1 GI:3668239
 KEYWORDS STS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 586)
 AUTHORS Myers, R.M.
 TITLE Human STSs (1998)

JOURNAL Unpublished (1998)
 COMMENT

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689

Email: myers@shgc.stanford.edu
 Primer A: AGAGACACGGTCAGCCCAAT
 Primer B: GCCAACTCTGAGTCATCCC
 STS size: 188
 PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Amplitaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs derived from W73140 -- Unigene.

FEATURES
 source

1. .586
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="19"
 /clone_lib="Human"

STS
 primer_bind 167..354
 primer_bind 167..186
 complement (335..354)
 ORIGIN

Query Match 30.7%; Score 482; DB 11; Length 586;
 Best Local Similarity 100.0%; Pred. No. 3.8e-131;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCTGG 1149
 DB 501 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCTGG 442
 QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
 DB 441 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 382
 QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269
 DB 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 322
 QY 1270 CCGGCATCCCACTGTCTCCAGGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC 1329
 DB 321 CCGGCATCCCACTGTCTCCAGGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC 262
 QY 1330 CCAGAGATGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 1389
 DB 261 CCAGAGATGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 202
 QY 1390 GTCTGCTTCCCACTGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 1449
 DB 201 GTCTGCTTCCCACTGTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGAGTCTCAG 142


```
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509
Db 141 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 82

QY 1510 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGAGA 1569
Db 81 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGAGA 22

QY 1570 AG 1571
Db 21 AG 20

RESULT 12
AX429955/c
LOCUS AX429955
DEFINITION Sequence 94 from Patent WO0198339.
ACCESSION AX429955
VERSION AX429955.1 GI:21541119
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: WO 0198339-A 94 27-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..735
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 3.9e-131; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0;

QY 1090 CAGGGTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1149
Db 500 CAGGGTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 441

QY 1150 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 440 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 381

QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269
Db 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 321

QY 1270 CCGGCATCCCACTGCTGTCAGGAGACCGCCTGACACTCTCTTTTCAGACCTCTATCTTC 1329
Db 320 CCGGCATCCCACTGCTGTCAGGAGACCGCCTGACACTCTCTTTTCAGACCTCTATCTTC 261

QY 1330 CCAGAGATTGTGAGATTGTTCTCTCTCAGCCCTGACCCCATGCTCTCTGACTCAGG 1389
Db 260 CCAGAGATTGTGAGATTGTTCTCTCTCAGCCCTGACCCCATGCTCTCTGACTCAGG 201

QY 1390 GTCTGCTTCCCACTGAGCGGTGACCGTGTCTCTCTAGTTGAACCTCTGGGAACAATTC 1449
Db 200 GTCTGCTTCCCACTGAGCGGTGACCGTGTCTCTCTAGTTGAACCTCTGGGAACAATTC 141

QY 1450 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509
Db 140 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 81

QY 1510 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGAGA 1569
Db 80 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGAGA 21
```

```
QY 1570 AG 1571
Db 20 AG 19

RESULT 13
BD139877/c
LOCUS BD139877
DEFINITION Compounds for immunotherapy and diagnosis of breast cancer and
methods for their use.
ACCESSION BD139877
VERSION BD139877.1 GI:23234822
KEYWORDS JP 2002507387-A/94.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 738)
AUTHORS Reed,S.G. and Xu,J.
TITLE Compounds for immunotherapy and diagnosis of breast cancer and
methods for their use
JOURNAL Patent: JP 2002507387-A 94 12-MAR-2002;
CORIXA CORP
COMMENT OS Homo sapiens (human)
PN JP 2002507387-A/94
PD 12-MAR-2002
PP 22-DEC-1998 JP 2000526543
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN
G REED,JIANGCHUN XU
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
PC C07K14/47,
PC
C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC
53,
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00
CC Compounds for immunotherapy and diagnosis
of breast cancer and
CC their use
CC Location/Qualifiers
FH Key 1..738
FT source /organism="Homo sapiens (human)"
FEATURES
source
1..738
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 738;
Best Local Similarity 100.0%; Pred. No. 3.9e-131; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0;

QY 1090 CAGGGTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1149
Db 503 CAGGGTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 444

QY 1150 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 443 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 384

QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269
Db 383 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 324

QY 1270 CCGGCATCCCACTGCTGTCAGGAGACCGCCTGACACTCTCTTTTCAGACCTCTATCTTC 1329
Db 323 CCGGCATCCCACTGCTGTCAGGAGACCGCCTGACACTCTCTTTTCAGACCTCTATCTTC 264

QY 1330 CCAGAGATTGTGAGATTGTTCTCTCTCAGCCCTGACCCCATGCTCTCTGACTCAGG 1389
Db 1330 CCAGAGATTGTGAGATTGTTCTCTCTCAGCCCTGACCCCATGCTCTCTGACTCAGG 1389
```


Db 263 CCAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCATGTCTCTCGGACTCAGG 204
Qy 1390 GTCTGCTTCCCCACATTTGGGTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449
Db 203 GTCTGCTTCCCCACATTTGGGTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 144
Qy 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACATTTTCATCTCAAGC 1509
Db 143 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACATTTTCATCTCAAGC 84
Qy 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569
Db 83 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 24
Qy 1570 AG 1571
Db 23 AG 22
RESULT 14
LOCUS CQ720600 1143 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 6534 from Patent WO02068579.
ACCESSION CQ720600
VERSION CQ720600.1 GI:42281457
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
PATENT: WO 02068579-A 6534 06-SEP-2002;
PE CORPORATION (NY) (US)
FEATURES
source
1. .1143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 1143;
Best Local Similarity 100.0%; Pred. No. 4e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGTGATTTCTGGGGGCGCTGTGTTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 1149
Db 657 CAGGGTGATTTCTGGGGGCGCTGTGTTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 716
Qy 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 717 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 776
Qy 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1269
Db 777 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 836
Qy 1270 CCGGCATCCCACTGTGTGAGGAGACGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1329
Db 837 CCGGCATCCCACTGTGTGAGGAGACGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 896
Qy 1330 CCAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCTCTGGACTCAGG 1389
Db 897 CCAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCTCTGGACTCAGG 956
Qy 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449
Db 957 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1016
Qy 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509

Db 1017 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1076
Qy 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569
Db 1077 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1136
Qy 1570 AG 1571
Db 1137 AG 1138
RESULT 15
LOCUS AY461805 1301 bp mRNA linear PRI 03-DEC-2003
DEFINITION Homo sapiens kallikrein 5 isoform 3 preproprotein (KLK5) mRNA,
complete cds, alternatively spliced.
ACCESSION AY461805
VERSION AY461805.1 GI:38564772
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL Michael, I.P., Kuriender, L., Du, D.C. and Diamandis, E.P.
Cloning of new splice variants of the human kallikrein gene 5
Unpublished
REFERENCE 2 (bases 1 to 1301)
AUTHORS Michael, I.P., Kuriender, L., Du, D.C. and Diamandis, E.P.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
FEATURES
source
1. .1301
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. .1301
/gene="KLK5"
12. .626
/genes="KLK5"
/note="alternatively spliced"
/codon_start=1
/product="kallikrein 5 isoform 3 preproprotein"
/protein_id="AA23814.1"
/db_xref="GI:38564773"
GSDILGAGAGEDARDSDSSRIINGSDDMHTQWQAALLRPQLYCGAVLVHPQW
LTAACHCKKVPVRVRLGHYSLSVPYSGGQMFQGVKSIHPHGYSHPGHSDNLMILKN
RRIRPTKDVPRPINVSSHCPSPAGTKCLVSGWGTTKSPQGBCPGSS"
ORIGIN
Query Match 30.7%; Score 482; DB 9; Length 1301;
Best Local Similarity 100.0%; Pred. No. 4.1e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGTGATTTCTGGGGGCGCTGTGTTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 1149
Db 820 CAGGGTGATTTCTGGGGGCGCTGTGTTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 879
Qy 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 880 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 939
Qy 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1269
Db 940 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 999
Qy 1270 CCGGCATCCCACTGTGTGAGGAGACGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1329

```
Db      ||||| 1000 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTGAGACCCCTCATTCCTTC 1059
Qy      ||||| 1330 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1389
Db      ||||| 1060 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1119
Qy      ||||| 1390 GTCTGCTTCCCCACATGCGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
Db      ||||| 1120 GTCTGCTTCCCCACATGCGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1179
Qy      ||||| 1450 CAAAACTGTCCAGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGC 1509
Db      ||||| 1180 CAAAACTGTCCAGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGC 1239
Qy      ||||| 1510 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569
Db      ||||| 1240 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1299
Qy      || 1570 AG 1571
Db      || 1300 AG 1301
```

Search completed: July 4, 2005, 09:55:43
Job time : 4534.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 552.864 Seconds
(without alignments)
16821.327 Million cell updates/sec

Title: US-09-936-271c-13_COPY_10000_11570
Perfect score: 1571
Sequence: 1 agggagtgatgggaattga.....cccgaaataaactgagaag 1571

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseqn_16Dec04:*
1: Geneseqn_1980s:*
2: Geneseqn_1980s:*
3: Geneseqn_2000s:*
4: Geneseqn_2001as:*
5: Geneseqn_2001bs:*
6: Geneseqn_2002as:*
7: Geneseqn_2002bs:*
8: Geneseqn_2003as:*
9: Geneseqn_2003bs:*
10: Geneseqn_2003cs:*
11: Geneseqn_2003ds:*
12: Geneseqn_2004as:*
13: Geneseqn_2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	11570	3	AAA95905
2	1571	100.0	11570	12	AdK52482 Human KLK
3	1571	100.0	11570	13	AdR72623 Human ren
4	1571	100.0	11570	13	AdR72875 Human ova
5	541	34.4	567	12	ACH74409 Human gen
C 6	482	30.7	586	6	ABL63579 Breast ca
C 7	482	30.7	586	6	ABL63996 Breast ca
C 8	482	30.7	735	3	AACT9469 CDNA sequ
C 9	482	30.7	735	6	ABK29013 Human bre
C 10	482	30.7	738	2	AAK84240 DNA encod
11	482	30.7	1260	10	AdB80489 Ovarian c
12	482	30.7	1260	11	AdN39197 Cancer/an
13	482	30.7	1260	13	AdR25550 Breast ca
14	482	30.7	1302	2	AAK57990
15	482	30.7	1302	2	AAK57989 Human BS2
16	482	30.7	1381	2	AAZ06259 Human sec
17	482	30.7	1381	8	ADA40370 Human sec
18	482	30.7	1381	10	ADA56535 Gene enco
19	482	30.7	1387	13	AdR72622 Human ren
20	482	30.7	1387	13	AdR72874 Human ova

21	482	30.7	1439	2	AAZ06260
22	482	30.7	1439	8	ADA40371
23	482	30.7	1439	10	ADA56536 Gene enco
24	482	30.7	1441	8	ADA40368
25	482	30.7	1441	10	ADA56533 Gene enco
26	482	30.7	1481	12	ADQ22212
27	482	30.7	1490	8	ADA40367
28	482	30.7	1490	10	ADA56532
29	482	30.7	1499	2	AAK60578
30	482	30.7	1516	2	AAZ06245
31	482	30.7	1516	8	ADA40369
32	482	30.7	1516	10	ADA56534
33	482	30.7	1570	3	AAZ65070
34	482	30.7	1570	4	AAZ21471
35	482	30.7	1570	5	AAF44216
36	482	30.7	1570	6	ABK28605
37	482	30.7	1570	8	ACA64363
38	482	30.7	1570	8	ACA03830
39	482	30.7	1570	8	ABX89368
40	482	30.7	1570	8	ABX80822
41	482	30.7	1570	8	ACD44331
42	482	30.7	1570	8	ACD42022
43	482	30.7	1570	8	ABX79502
44	482	30.7	1570	8	ACA93523
45	482	30.7	1570	8	ABX81205

ALIGNMENTS

RESULT 1
AAA95905
ID AAA95905 standard; DNA; 11570 BP.
XX
AC AAA95905;
XX
DT 02-FEB-2001 (first entry)
XX
DE Human KLK-L2 gene.
XX
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KW kallikrein-like protein; serine protease; cytosolic; cancer;
KW prostrate cancer; ds.
XX
OS Homo sapiens.
XX
PN WO200053776-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-CA000258.
XX
PR 11-MAR-1999; 99US-0124260P.
PR 01-APR-1999; 99US-0127386P.
PR 21-JUL-1999; 99US-0144919P.
XX
PA (MOUN) MOUNT SINAI HOSPITAL.
XX
PI Yousef GM, Diamandis EP;
XX
DR WPI; 2000-587440/55.
DR P-PSDB; AAB21296.
XX
PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
XX protein mediated disorders, especially cancer.
PS Claim 1; Page 143-149; 184pp; English.
XX
CC The present sequence is the coding sequence of the human KLK-L2 gene,
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like
CC proteins are a subgroup of the serine protease enzyme family. They
CC catalyse the selective cleavage of specific polypeptide precursors to
CC release peptides with potent biological activity. Nucleic acids encoding

XX Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.
DE kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;
KW cancer metastasis; chemotherapy; human; serine protease;
KW chromosome 19q13.4; KLK5; ds; gene.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 2221..11247
FT /*tag= b
FT /product= "Human renal cell carcinoma-related kallikrein
FT 5 (hK5) protein"
FT exon 2221..2293
FT /*tag= a
FT /number= 1
FT intron 2294..4761
FT /*tag= c
FT /number= 1
FT exon 4762..5023
FT /*tag= d
FT /number= 2
FT intron 5024..5762
FT /*tag= e
FT /number= 2
FT exon 5763..6019
FT /*tag= f
FT /number= 3
FT intron 6020..6104
FT /*tag= g
FT /number= 3
FT exon 6105..6238
FT /*tag= h
FT /number= 4
FT exon 6239..11091
FT /*tag= i
FT /number= 4
FT exon 11092..11247
FT /*tag= j
FT /number= 5
XX WO2004077060-A2.
PN 10-SEP-2004.
XX
XX 26-FEB-2004; 2004WO-CA000280.
XX
XX 27-FEB-2003; 2003US-0451382P.
XX (MOUN) MOUNT SINAI HOSPITAL.
XX Diamandis EP, Petraki CD;
PI WPI; 2004-662077/64.
XX P-PSDB; ADR72621.
XX
XX Detecting kallikrein polypeptides/polynucleotides associated with renal
XX cell carcinoma in a patient, for diagnosing/treating the disease,
XX comprises detecting /identifying kallikrein polypeptides/polynucleotides
XX in a sample.
XX
XX Example 1; SEQ ID NO 3; 53pp; English.
XX
XX The invention relates to a novel method for detecting kallikrein
XX polypeptides, or the polynucleotides encoding them, associated with renal
XX cell carcinoma. The method comprises obtaining a sample from a patient
XX and detecting kallikrein polypeptides, or their encoding polynucleotides,
XX where the kallikrein polypeptides are selected from the group consisting
XX of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The
XX detected amounts of the kallikrein polypeptides are compared to standard
XX amounts. The molecules of the invention demonstrate cytostatic activity
XX whilst the methods and kit may be useful for detecting, characterising,

CC preventing and treating renal cell carcinoma. Furthermore, the methods
CC may be useful for evaluating the probability of the presence of malignant
CC or pre-malignant cells and for detecting and quantitating tumour growth
CC and cancer metastasis. Finally, the methods may be utilised to confirm
CC the absence or removal of all tumour tissue following surgery, cancer
CC chemotherapy and/or radiation therapy and to monitor cancer chemotherapy
CC and tumour reappearance. The current sequence is that of the human
CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the
CC invention which encodes a secreted serine protease and is located at
CC chromosome 19q13.4.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 1571; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGAGGTATGGGAAATTGAAGACAGGAAACACAAATTAAGTCCAAAGCAATGGATTCTTA 60
Db 10000 AGGGAGGTATGGGAAATTGAAGACAGGAAACACAAATTAAGTCCAAAGCAATGGATTCTTA 10059
Qy 61 TTGGGAGTGTCTGCCCCCTAGAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA 120
Db 10060 TTGGGAGTGTCTGCCCCCTAGAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA 10119
Qy 121 CAACTATATGGAGGGGCATTACTGGCAACTAATGGATAGATGCAAGTGTCTGTTCAC 180
Db 10120 CAACTATATGGAGGGGCATTACTGGCAACTAATGGATAGATGCAAGTGTCTGTTCAC 10179
Qy 181 ATGCTATATGACACACGGCAGGCTCCACAAACCAATTTATCCAGTTCAGATGCCAC 240
Db 10180 ATGCTATATGACACACGGCAGGCTCCACAAACCAATTTATCCAGTTCAGATGCCAC 10239
Qy 241 AGTGCCAGATCGAGGAACCTTCATCCAGGGCTGAGAACCGTATTTTTCAGAGGGAG 300
Db 10240 AGTGCCAGATCGAGGAACCTTCATCCAGGGCTGAGAACCGTATTTTTCAGAGGGAG 10299
Qy 301 GTATAAGGATGGGTGTGGAGAAATGGGAAGCAAGTGTGTCTCAGTAAGAGAAATAA 360
Db 10300 GTATAAGGATGGGTGTGGAGAAATGGGAAGCAAGTGTGTCTCAGTAAGAGAAATAA 10359
Qy 361 GGCCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGGAGGGAGAGATACACGATGA 420
Db 10360 GGCCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGGAGGGAGAGATACACGATGA 10419
Qy 421 GGGAGACAGGCTGGAAACAGAAAGTAGACAGCAAGATTTCAGATCTCGAGAGAGGGTCA 480
Db 10420 GGGAGACAGGCTGGAAACAGAAAGTAGACAGCAAGATTTCAGATCTCGAGAGAGGGTCA 10479
Qy 481 CAGACCCCCCGAAATGATGTGTGGACAAACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 540
Db 10480 CAGACCCCCCGAAATGATGTGTGGACAAACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 10539
Qy 541 AGTGACAAATGGGGTCTAAAGTTGAACCTTGGAGCCAGGCATGGTGGCTCAGCCCTGTA 600
Db 10540 AGTGACAAATGGGGTCTAAAGTTGAACCTTGGAGCCAGGCATGGTGGCTCAGCCCTGTA 10599
Qy 601 ATCCCAACACTTTGGAGGCTGAGGTGGGGCAATCATTGAGGCCAGGAGTTCAGACCAG 660
Db 10600 ATCCCAACACTTTGGAGGCTGAGGTGGGGCAATCATTGAGGCCAGGAGTTCAGACCAG 10659
Qy 661 CTGGGCCAATCGTGTGAAACCCCGTCTCTACAAAAAAATACAAAAATTTAGCGGGTGT 720
Db 10660 CTGGGCCAATCGTGTGAAACCCCGTCTCTACAAAAAAATACAAAAATTTAGCGGGTGT 10719
Qy 721 GGTGATGGACACCTGTAGTCACAGCTACTTTGGAGGCTGAGGCGAGGAGAAATGCTTGAAC 780
Db 10720 GGTGATGGACACCTGTAGTCACAGCTACTTTGGAGGCTGAGGCGAGGAGAAATGCTTGAAC 10779
Qy 781 CCGGAGATGGAGGCTGCGAGTGAGCTGAGGTGAGGCCACTGCGCTCCAACTGGGCAACA 840
Db 10780 CCGGAGATGGAGGCTGCGAGTGAGGTGAGGCCACTGCGCTCCAACTGGGCAACA 10839

[illegible]

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above), where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, a method of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subexpression, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

Sequence 567 BP; 105 A; 197 C; 125 G; 140 T; 0 U; 0 Other;

Query Match 34.4%; Score 541; DB 12; Length 567;
Best Local Similarity 100.0%; Pred. No. 7e-128;
Matches 541; Conservative 0; Mismatches 0; Indels

Qy	1031	GTCTCCTTCTGCCCCACATTTTGCCACATCTCTGCGCTCTCTCATGCCCCCCTTTCTCTCTCTGC	1090
Db	1	GTCTCCTTCTGCCCCACATTTTGCCACATCTCTGCGCTCTCTCATGCCCCCCTTTCTCTCTCTGC	60
Qy	1091	AGGCTGATTTCTGGGGGGCCTGTGCTGTGCAATGCTCCCTCGAGGGACTCGTGTCTGCGG	1150
Db	61	AGGCTGATTTCTGGGGGGCCTGTGCTGTGCAATGCTCCCTCGAGGGACTCGTGTCTCTGCGG	120
Qy	1151	GAGATTACCTTTGTCCCGGCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTCA	1210
Db	121	GAGATTACCTTTGTCCCGGCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTCA	180
Qy	1211	CCAGTGGATCCAGGAACCATCAGGCCAATCTCTGAGTCATCCAGGACTCAGCAC	1270
Db	181	CCAGTGGATCCAGGAACCATCAGGCCAATCTCTGAGTCATCCAGGACTCAGCAC	240
Qy	1271	CGGCATCCCCACCTGTCTGAGGGACAGCCCTGACATCTTTCAGACCCCTCATTTCTTCC	1330
Db	241	CGGCATCCCCACCTGTCTGAGGGACAGCCCTGACATCTTTCAGACCCCTCATTTCTTCC	300
Qy	1331	CAGAGATGTTGAGAAATGTTCACTCTCAGGCCCTGACCCCATGTCTCTGGACTCAGGG	1390
Db	301	CAGAGATGTTGAGAAATGTTCACTCTCTCAGGCCCTGACCCCATGTCTCTGGACTCAGGG	360
Qy	1391	TCTCTTCCCCCACAATTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTTCC	1450
Db	361	TCTCTTCCCCCACAATTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTTCC	420
Qy	1451	AAAACTGTCCAGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGCT	1510
Db	421	AAAACTGTCCAGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGCT	480

PR 01-NOV-2000; 2000US-0244867P.
XX 01-NOV-2000; 2000US-0245084P.
PA (AVAL-) AVALON PHARM.
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 1916; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX tumour
XX
SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;

Query Match 30.7%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 8.3e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCG 1149
DB |||||
QY 501 CAGGGTGAATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCG 442
DB |||||
QY 1150 GGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 1209
DB |||||
QY 441 GGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 382
DB |||||
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTTCATCCAGGACTCAGCACA 1269
DB |||||
QY 381 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTTCATCCAGGACTCAGCACA 322
DB |||||
QY 1270 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1329
DB |||||
QY 321 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 262
DB |||||
QY 1330 CCAGAGATGTTGAGAAATTTCTCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 1389
DB |||||
QY 261 CCAGAGATGTTGAGAAATTTCTCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 202
DB |||||
QY 1390 GTCTGTCTCCCAATGTTGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
DB |||||
QY 201 GTCTGTCTCCCAATGTTGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 142
DB |||||
QY 1450 CAAAACGTCTCAGGCGCGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
DB |||||
QY 141 CAAAACGTCTCAGGCGCGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 82
DB |||||
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1569
DB |||||
QY 81 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 22
DB |||||
QY 1570 AG 1571

Db 21 AG 20

RESULT 7

ABL63996/c
ID ABL63996 standard; DNA; 586 BP.
XX
XX ABL63996;
XX
XX 15-MAY-2002 (first entry)
DT
XX Breast cancer related gene sequence SEQ ID NO:2333.
DE
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
XX Homo sapiens.
PN WO200194629-A2.
XX
XX 13-DEC-2001.
PD
XX 30-MAY-2001; 2001WO-US010838.
PF
XX 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
PA
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI

PI Soppet DR, Weaver Z;
XX WPI: 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX
PS Claim 1; SEQ ID NO 2333; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
XX tumour
SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;
Query Match 30.7%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 8.3e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGATTCCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 1149
DB 501 CAGGGTGATTCCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 442
QY 1150 GGAGATTACCTTGTGTCGGGCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 441 GGAGATTACCTTGTGTCGGGCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 1269
DB 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 322
QY 1270 CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC 1329
DB 321 CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC 262
QY 1330 CCAGAGATGTTGAGATGTTTATCTCTCTCCAGCCCTGACCCCAATGTCTCTGGACTCAGG 1389
DB 261 CCAGAGATGTTGAGATGTTTATCTCTCTCCAGCCCTGACCCCAATGTCTCTGGACTCAGG 202
QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGAAACAATTC 1449
DB 201 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGAAACAATTC 142
QY 1450 CAAACATGTCAGGGCGGGGTTCGGTCTCAATCTCTCCCTGGGCACTTTCATCTCTCAAGC 1509
DB 141 CAAACATGTCAGGGCGGGGTTCGGTCTCAATCTCTCCCTGGGCACTTTCATCTCTCAAGC 82
QY 1510 TCAGGGCCCACTCCCTCTCTGACGCTCTGACCCCAAAATTTAGTCCAGAAATAAAGTGA 1569
DB 81 TCAGGGCCCACTCCCTCTCTGACGCTCTGACCCCAAAATTTAGTCCAGAAATAAAGTGA 22
QY 1570 AG 1571
DB 21 AG 20
RESULT 8
AAC79469/c

AAC79469 standard; cDNA; 735 BP.
AAC79469;
07-FEB-2001 (first entry)
cDNA sequence of human breast tumour clone B541S.
Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;
vaccine; ss.
Homo sapiens.
W0200061756-A2.
19-OCT-2000.
10-APR-2000; 2000MO-US009688.
09-APR-1999; 99US-00288950.
02-JUL-1999; 99US-00346327.
(CORI-) CORIXA CORP.
Reed SG, Xu J, Dillon DC;
WPI: 2000-638568/61.
A novel isolated polypeptide comprising an immunogenic portion of a
breast cancer protein useful in the detection and treatment of breast
cancer.
Claim 4; Page 88; 95pp; English.
The present sequence was isolated from a breast tumour cDNA library. It
is provided in a specification relating to compounds for immunotherapy
and diagnosis of breast cancer. Breast tumour antigens and the
polynucleotides that encode them may be used in the production of a
pharmaceutical composition to be used in the treatment of breast cancer.
Proliferated T cells and incubated antigen presenting cells are also
required. The polypeptides and polynucleotides may also be used to
produce a vaccine
Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 3; Length 735;
Best Local Similarity 100.0%; Pred. No. 9e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGATTCCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 1149
DB 500 CAGGGTGATTCCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 441
QY 1150 GGAGATTACCTTGTGTCGGGCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 440 GGAGATTACCTTGTGTCGGGCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 381
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 1269
DB 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 321
QY 1270 CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC 1399
DB 320 CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC 261
QY 1330 CCAGAGATGTTGAGATGTTTATCTCTCTCCAGCCCTGACCCCAATGTCTCTGGACTCAGG 1389
DB 260 CCAGAGATGTTGAGATGTTTATCTCTCTCCAGCCCTGACCCCAATGTCTCTGGACTCAGG 201
QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGAAACAATTC 1449
DB 200 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGAAACAATTC 141

```
QY 1450 CAAAACGTGTCAGGGCGGGGTTGCGTCTCAATCTCCTCGGGGACATTTTCATCTCAAGC 1509
Db 140 CAAAACGTGTCAGGGCGGGGTTGCGTCTCAATCTCCTCGGGGACATTTTCATCTCAAGC 81
QY 1510 TCAGGGCCCATCCCTTCTCTGCGACTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569
Db 80 TCAGGGCCCATCCCTTCTCTGCGACTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 21
QY 1570 AG 1571
Db 20 AG 19

RESULT 9
ABK29013/c
ID ABK29013 standard; cDNA; 735 BP.
XX
AC ABK29013;
XX
DT 23-APR-2002 (first entry)
XX
XX Human breast tumour polypeptide cDNA clone #42.
XX
DE Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW immunostimulant.
XX
OS Homo sapiens.
XX
PN W0200198339-A2.
XX
PD 27-DEC-2001.
XX
XX 12-JUN-2001; 2001WO-US019032.
XX
XX 22-JUN-2000; 2000US-00602877.
PR 12-OCT-2000; 2000US-00687507.
PR 06-FEB-2001; 2001US-00778381.
XX
XX (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;
XX
XX WPI; 2002-147792/19.
XX
XX Polynucleotides encoding breast tumor polypeptides, useful for treating
PT breast cancer or stimulating an immune response.
XX
XX Claim 1; Page 140; 150pp; English.
XX
XX The invention relates to polynucleotides encoding breast tumour
CC polypeptides. The sequences are useful for treating cancer, preferably
CC breast cancer, in a patient or for stimulating an immune response. The
CC polynucleotides and polypeptides are also useful in the diagnosis and
CC monitoring of breast cancer. A method for detecting the presence of a
CC cancer in a patient, comprises obtaining a biological sample from the
CC patient, contacting the biological sample with a binding agent that binds
CC to a breast tumour polypeptide, detecting in the sample an amount of
CC polypeptide that binds to the binding agent, and comparing the amount of
CC polypeptide to a predetermined cut-off value, therefore determining the
CC presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC represent cDNA clones encoding human breast tumour polypeptides of the
CC invention
XX
SQ Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 9e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTTCTGGGGGCGCTGCTGCTGCAATGGCTCCCTGCGAGGACATCGTGCTCTGG 1149
Db 500 CAGGGTGATTTCTGGGGGCGCTGCTGCTGCAATGGCTCCCTGCGAGGACATCGTGCTCTGG 441
```

```
QY 1150 GGAGATTACCTTGTGTCGCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTCAAGTTC 1209
Db 440 GGAGATTACCTTGTGTCGCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTCAAGTTC 381
QY 1210 ACCAAGTGGATCCAGGAAACCAATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA 1269
Db 380 ACCAAGTGGATCCAGGAAACCAATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA 321
QY 1270 CCGGCATCCCACTCTGTCAGGGCAGCCCTGACACTCCTTTTCAGACCCCTCATTTCTTTC 1329
Db 320 CCGGCATCCCACTCTGTCAGGGCAGCCCTGACACTCCTTTTCAGACCCCTCATTTCTTTC 261
QY 1330 CCAGAGATGTGAGAAATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTCGGACTCAGG 1389
Db 260 CCAGAGATGTGAGAAATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTCGGACTCAGG 201
QY 1390 GTCTGCTTCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTTTC 1449
Db 200 GTCTGCTTCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTTTC 141
QY 1450 CAAAACGTGTCAGGGCGGGGTTGCGTCTCAATCTCCTCGGGGACATTTTCATCTCAAGC 1509
Db 140 CAAAACGTGTCAGGGCGGGGTTGCGTCTCAATCTCCTCGGGGACATTTTCATCTCAAGC 81
QY 1510 TCAGGGCCCATCCCTTCTCTGCGACTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569
Db 80 TCAGGGCCCATCCCTTCTCTGCGACTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 21
QY 1570 AG 1571
Db 20 AG 19

RESULT 10
AXX84240/c
ID AAX84240 standard; cDNA; 738 BP.
XX
AC AAX84240;
XX
DT 08-SEP-1999 (first entry)
XX
DE DNA encoding human breast tumour protein immunogenic fragment.
XX
KW Breast tumour protein; immunogenic fragment; vaccine; detection;
KW breast cancer development; therapy; ss.
XX
OS Homo sapiens.
XX
PN W09933869-A2.
XX
XX 08-JUL-1999.
XX
XX 22-DEC-1998; 98WO-US027416.
XX
XX 24-DEC-1997; 97US-00998253.
PR 24-DEC-1997; 97US-00998255.
PR 17-JUL-1998; 98US-00118554.
PR 17-JUL-1998; 98US-00118627.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Xu J;
XX
XX WPI; 1999-405486/34.
XX
XX New breast tumor protein genes used, in vaccines for immunotherapy, or
PT for diagnosis of breast cancer.
XX
XX Claim 3; Page 70; 70pp; English.
XX
XX This sequence encodes a human breast tumour protein immunogenic fragment
CC of the invention. The polypeptides or nucleic acids encoding them are
CC useful in vaccines and pharmaceutical compositions for manufacture of
```

CC medications for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 738 BP; 161 A; 179 C; 226 G; 172 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 2; Length 738;

Best Local Similarity 100.0%; Pred. No. 9.1e-113; Indels 0; Gaps 0;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGAGCTCGTCTCTGG 1149

DB 503 CAGGGTGAATTCGTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGAGCTCGTCTCTGG 444

QY 1150 GAGATTAACCTTGTGCCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209

DB 443 GAGATTAACCTTGTGCCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 384

QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCCTGAGTTCATCCAGGACTCAGCACA 1269

DB 383 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCCTGAGTTCATCCAGGACTCAGCACA 324

QY 1270 CCGGCATCCCACTGCTGAGGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTCCTTC 1329

DB 323 CCGGCATCCCACTGCTGAGGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTCCTTC 264

QY 1330 CCAGAGATGTGAGATGTTTCATCTCTCAGCCCTGACCCCAATGTCTCTGGACTCAGG 1389

DB 263 CCAGAGATGTGAGATGTTTCATCTCTCAGCCCTGACCCCAATGTCTCTGGACTCAGG 204

QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCGGACAAATTC 1449

DB 203 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCGGACAAATTC 144

QY 1450 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 1509

DB 143 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 84

QY 1510 TCAGGGCCCAATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1569

DB 83 TCAGGGCCCAATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 24

QY 1570 AG 1571

DB 23 AG 22

RESULT 11

ADB80489

ID ADB80489 standard; DNA; 1260 BP.

XX AC

XX ADB80489;

XX AC

DT 04-DEC-2003 (first entry)

XX

DE Ovarian cancer-associated transcript #27.

XX

KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;

KW post-operative chemotherapy; radiation therapy; tumour prognosis;

XX pre-cancerous lesion detection; ds; gene.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX CDS 43..924

FT /*cag= a

FT

XX

PN

XX

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

W02002102235-A2.

27-DEC-2002.

18-JUN-2002; 2002WO-US019297.

18-JUN-2001; 2001US-0299234P.

27-AUG-2001; 2001US-0315287P.

05-SEP-2001; 2001US-0317544P.

13-NOV-2001; 2001US-0350666P.

12-APR-2002; 2002US-0372246P.

(E0SB-) EOS BIOTECHNOLOGY INC.

Mack DH, Gish KC;

WPI; 2003-167431/16.

P-PSDB; ADB80490.

Detecting an ovarian cancer-associated transcript in a cell from a

patient, comprises contacting a biological sample from the patient with a

polynucleotide that hybridizes to an ovarian cancer gene.

Claim 10; Page 292; 332pp; English.

The invention relates to a method of detecting an ovarian cancer-

associated transcript in a cell from a patient, by contacting a

biological sample from the patient with a polynucleotide that selectively

hybridizes to a sequence at least 80% identical to any of one of 80

nucleic acid sequences given in the specification. The method is useful

in diagnosing ovarian cancer and in identifying and using agents and/or

targets that inhibit ovarian cancer. The nucleic acid molecule,

polypeptide and the antibody may also be used in detecting ovarian

cancers, monitoring and early detection of relapse following treatment,

monitoring response to therapy, selecting patients for post-operative

chemotherapy or radiation therapy, in selecting mode of therapy,

determining tumour prognosis, early detection of pre-cancerous lesions,

and as vaccines. This sequence corresponds to one of the nucleic acids

used for the detection method of the invention.

SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 10; Length 1260;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGAGCTCGTCTCTGG 1149

DB 766 CAGGGTGAATTCGTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGAGCTCGTCTCTGG 825

QY 1150 GAGATTAACCTTGTGCCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209

DB 826 GAGATTAACCTTGTGCCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 885

QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCCTGAGTTCATCCAGGACTCAGCACA 1269

DB 886 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCCTGAGTTCATCCAGGACTCAGCACA 945

QY 1270 CCGGCATCCCACTGCTGAGGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTCCTTC 1329

DB 946 CCGGCATCCCACTGCTGAGGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTCCTTC 1005

QY 1330 CCAGAGATGTGAGATGTTTCATCTCTCAGCCCTGACCCCAATGTCTCTGGACTCAGG 1389

DB 1006 CCAGAGATGTGAGATGTTTCATCTCTCAGCCCTGACCCCAATGTCTCTGGACTCAGG 1065

QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCGGACAAATTC 1449

DB 1066 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCGGACAAATTC 1125

QY 1450 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 1509

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

DB

Db 1126 CAAAACTGTCCAGGGCGGGGTTCGGTCTCAATCTCCTCGGGGCACTTTTCATCTCTCAAGC 1185
Qy 1510 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1569
Db 1186 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1245
Qy 1570 AG 1571
Db 1246 AG 1247
RESULT 12
ADN39197
ID ADN39197 standard; cDNA; 1260 BP.
XX
AC ADN39197;
XX
DT 17-JUN-2004 (first entry)
XX
XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:515.
XX
XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
XX
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-035394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR P-PSDB; ADN39198.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 515; 1385pp; English.
PS
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acid, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 11; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.1e-112;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGCTCGG 1149
Db 766 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGCTCGG 825
Qy 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209
Db 826 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 885
Qy 1210 ACCAAGTGATCCAGGAACCATTCAGGGCAACTCTCTAGTCTATCCAGGACTCAGACA 1269
Db 886 ACCAAGTGATCCAGGAACCATTCAGGGCAACTCTCTAGTCTATCCAGGACTCAGACA 945
Qy 1270 CCGGCATCCCACTCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1329
Db 946 CCGGCATCCCACTCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1005
Qy 1330 CCAGAGATGTTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGGACTCAGG 1389
Db 1006 CCAGAGATGTTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCTCTGGACTCAGG 1065
Qy 1390 GTCTGTTCCCCACATTCGGGCTGACCGTGTCTCTTAGTTGAACCTCGGGAACAATTC 1449
Db 1066 GTCTGTTCCCCACATTCGGGCTGACCGTGTCTCTTAGTTGAACCTCGGGAACAATTC 1125
Qy 1450 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
Db 1126 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1185
Qy 1510 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1569
Db 1186 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1245
Qy 1570 AG 1571
Db 1246 AG 1247
RESULT 13
ADR25550
ID ADR25550 standard; DNA; 1260 BP.
XX
XX ADR25550;
XX
XX 21-OCT-2004 (first entry)
DT Breast cancer prognosis marker #1411.
XX
XX ds; breast cancer; prognosis; gene expression; diagnosis.
XX
XX Homo sapiens.
OS
XX
XX WO2004065545-A2.
PN

XX PD 05-AUG-2004.
XX PF 15-JAN-2004; 2004WO-US001100.
XX PR 15-JAN-2003; 2003US-00342887.
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX PA (NECA-) NETHERLANDS CANCER INST.
XX PI Van't Veer LJ, He Y;
XX DR WPI; 2004-593473/57.
XX PT Classifying a breast cancer patient according to prognosis comprises
XX PT determining the similarity between the level of expression of each of
XX PT five genes in a cell sample taken from patient, to control levels.
XX PS Disclosure; SEQ ID NO 1411; 226pp; English.
XX CC The invention relates to a method of classifying a breast cancer patient
XX CC according to prognosis by determining the similarity between the level of
XX CC expression of each of five genes for which markers are listed in the
XX CC specification, in a cell sample taken from the breast cancer patient, to
XX CC control levels of expression for each respective five genes to obtain a
XX CC patient similarity value. The methods are useful for classifying a breast
XX CC cancer patient according to prognosis. Kits and computer program products
XX CC are useful for data analysis using the diagnostic, prognostic and
XX CC statistical methods of the invention. This sequence corresponds to a
XX CC marker used in the method of the invention.
XX SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 13; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.1e-112;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGCTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149
DB 766 CAGGCTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 825
QY 1150 GGAGATTACCTCTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
DB 826 GGAGATTACCTCTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 885
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 1269
DB 886 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 945
QY 1270 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329
DB 946 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1005
QY 1330 CCAGAGATGTTGAGAATGTTATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
DB 1006 CCAGAGATGTTGAGAATGTTATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1065
QY 1390 GTCTGCTTCCCACTGAGCTGACCGTGTCTCTAGTTGAACCTCTGGGAACTTC 1449
DB 1066 GTCTGCTTCCCACTGAGCTGACCGTGTCTCTAGTTGAACCTCTGGGAACTTC 1125
QY 1450 CAAAACCTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGGCACTTCATCTCAAGC 1509
DB 1126 CAAAACCTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGGCACTTCATCTCAAGC 1185
QY 1510 TCAGGGCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAAGTGA 1569
DB 1186 TCAGGGCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAAGTGA 1245
QY 1570 AG 1571
DB 1246 AG 1247

RESULT 14

AA57990
ID AAX57990 standard; DNA; 1302 BP.
XX AC AAX57990;
XX DT 19-JUL-1999 (first entry)
XX DE Human BS247 specific polynucleotide #8.
XX KW BS247; detection; diagnosis; breast cancer; atypical hyperplasia;
XX KW fibroadenoma; cystic breast disease; gene therapy; ss.
XX OS Homo sapiens.
XX PN WO9922027-A1.
XX PD 06-MAY-1999.
XX PF 28-OCT-1998; 98WO-US022906.
XX PR 28-OCT-1997; 97US-0063431P.
XX PR 28-OCT-1997; 97US-00968838.
XX PA (ABBO) ABBOTT LAB.
XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
XX PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
XX PI Stroupe SD, Yu H;
XX DR WPI; 1999-312977/26.
XX CC Breast tissue derived cDNA contig and consensus polypeptide sequence.
XX Claim 1; Page 106; 112pp; English.
XX CC This sequence is a BS247 specific polynucleotide. The invention relates
XX CC to a method of detecting the presence of a target BS247 polynucleotide,
XX CC especially mRNA, in a test sample. BS247 polynucleotides are derived from
XX CC breast tissue. The polynucleotides, polypeptides or antibodies are useful
XX CC for providing information leading to the detection, diagnosis, staging,
XX CC monitoring, prognosis, in vivo imaging, prevention or treatment,
XX CC determining predisposition to, diseases and conditions of the breast,
XX CC such as breast cancer, atypical hyperplasia, fibroadenoma and cystic
XX CC breast disease. Drug treatment or gene therapy for breast cancer, can be
XX CC based on these identified gene sequences and the efficacy of any
XX CC particular therapy can be monitored. The BS247-derived reagents are
XX CC advantageous for detection of breast cancer due to their specificity. The
XX CC reagents also provide an alternative, non-surgical diagnostic method
XX CC capable of detecting early stage breast disease, such as cancer
XX SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 2; Length 1302;
Best Local Similarity 100.0%; Pred. No. 1.1e-112;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGCTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149
DB 811 CAGGCTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 870
QY 1150 GGAGATTACCTCTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
DB 871 GGAGATTACCTCTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 930
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 1269
DB 931 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 990
QY 1270 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329
DB 991 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1050

QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
|||
Db 1051 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1110
|||
QY 1390 GTCTGCTTCCCCACATGCGGTGACCGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1449
|||
Db 1111 GTCTGCTTCCCCACATGCGGTGACCGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1170
|||
QY 1450 CAAAACGTCTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
|||
Db 1171 CAAAACGTCTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1230
|||
QY 1510 TCAGGGCCCATCCCTCTCTGCAAGTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569
|||
Db 1231 TCAGGGCCCATCCCTCTCTGCAAGTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1290
|||
QY 1570 AG 1571
||
Db 1291 AG 1292

RESULT 15

AAX57989
ID AAX57989 standard; DNA; 1302 BP.

XX AAX57989;

XX 19-JUL-1999 (first entry)

XX Human BS247 specific polynucleotide #7.

XX BS247; detection; diagnosis; breast cancer; atypical hyperplasia;
XX fibroadenoma; cystic breast disease; gene therapy; ss.

XX Homo sapiens.

XX WO9922027-A1.

XX 06-MAY-1999.

XX 28-OCT-1998; 98WO-US022906.

XX 28-OCT-1997; 97US-0063431P.

XX 28-OCT-1997; 97US-00968838.

XX (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
XX Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
XX Stroupe SD, Yu H;

XX WPI; 1999-312977/26.

XX Breast tissue derived cDNA contig and consensus polypeptide sequence.

XX Claim 1; Page 105; 112pp; English.

XX This sequence is a BS247 specific polynucleotide. The invention relates
XX to a method of detecting the presence of a target BS247 polynucleotide,
XX especially mRNA, in a test sample. BS247 polynucleotides are derived from
XX breast tissue. The polynucleotides, polypeptides or antibodies are useful
XX for providing information leading to the detection, diagnosis, staging,
XX monitoring, prognosis, in vivo imaging, prevention or treatment,
XX determining predisposition to, diseases and conditions of the breast,
XX such as breast cancer, atypical hyperplasia, fibroadenoma and cystic
XX breast disease. Drug treatment or gene therapy for breast cancer, can be
XX based on these identified gene sequences and the efficacy of any
XX particular therapy can be monitored. The BS247-derived reagents are
XX advantageous for detection of breast cancer due to their specificity. The
XX reagents also provide an alternative, non-surgical diagnostic method
XX capable of detecting early stage breast disease, such as cancer

SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 2; Length 1302;
Best Local Similarity 100.0%; Pred. No. 1.le-112; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0;
QY 1090 CAGGGTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTCGG 1149
|||
Db 811 CAGGGTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTCGG 870
|||
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209
|||
Db 871 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 930
|||
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 1269
|||
Db 931 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 990
|||
QY 1270 CCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACATCTCTTTCAGACCCCTCATTCCTTC 1329
|||
Db 991 CCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACATCTCTTTCAGACCCCTCATTCCTTC 1050
|||
QY 1330 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
|||
Db 1051 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1110
|||
QY 1390 GTCTGCTTCCCCACATGCGGTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1449
|||
Db 1111 GTCTGCTTCCCCACATGCGGTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1170
|||
QY 1450 CAAAACGTCTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
|||
Db 1171 CAAAACGTCTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1230
|||
QY 1510 TCAGGGCCCATCCCTTCTCTGCAAGTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569
|||
Db 1231 TCAGGGCCCATCCCTTCTCTGCAAGTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1290
|||
QY 1570 AG 1571
||
Db 1291 AG 1292

Search completed: July 4, 2005, 03:03:35
Job time : 556.864 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:41:10 ; Search time 178.302 Seconds

(without alignments)
14417.061 Million cell updates/sec

Title: US-09-936-271C-13_COPY_10000_11570

Perfect score: 1571

Sequence: 1 agggagtgatgggaattga.....cccgagaataaactgagaag 1571

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1567.8	99.8	10818	4	US-09-949-016-13583
C 2	599	38.1	601	4	US-09-949-016-63416
C 3	482	30.7	735	3	US-09-602-877A-94
4	482	30.7	1143	4	US-09-949-016-1841
5	482	30.7	1499	4	US-09-509-908-1
C 6	480.6	30.6	601	4	US-09-949-016-63415
7	476	30.3	1476	2	US-08-824-874-2
8	476	30.3	1476	3	US-09-210-084-2
9	476	30.3	1476	4	US-09-784-762-2
10	460	29.3	1504	3	US-09-280-116-1
C 11	253	16.1	601	4	US-09-949-016-63414
C 12	223.4	14.2	68719	4	US-09-949-016-12799
13	223.4	14.2	68720	4	US-09-949-016-14296
14	222.8	14.2	32278	4	US-09-949-016-14575
15	222.6	14.2	176373	3	US-09-128-155-17
C 16	222.2	14.1	601	4	US-09-949-016-66914
C 17	222.2	14.1	177669	4	US-09-949-016-13713
18	221.8	14.1	80858	4	US-09-949-016-12659
19	221.8	14.1	80859	4	US-09-949-016-15715
C 20	220.8	14.1	69701	4	US-09-949-016-14187
C 21	220.8	14.1	73308	4	US-09-949-016-16326
C 22	219.8	14.0	300402	4	US-09-949-016-13632
C 23	219.4	14.0	38009	4	US-09-949-016-13617
C 24	219.4	14.0	70000	3	US-09-851-896-3
C 25	219.4	14.0	76399	4	US-09-949-016-16819
26	219.2	14.0	152331	3	US-09-128-155-16
27	218.8	13.9	71574	4	US-09-949-016-15580

28	218.6	13.9	25464	4	US-09-326-480A-4	Sequence 4, Appli
29	218.6	13.9	87039	4	US-09-949-016-15691	Sequence 15691, A
C 30	218.2	13.9	41755	4	US-09-949-016-15728	Sequence 15728, A
31	218	13.9	5695	4	US-09-949-016-13664	Sequence 13664, A
C 32	218	13.9	301828	4	US-09-949-016-13969	Sequence 13969, A
C 33	217.8	13.9	601	4	US-09-949-016-64323	Sequence 64323, A
C 34	217.8	13.9	36156	4	US-09-949-016-12128	Sequence 12128, A
C 35	217.8	13.9	36156	4	US-09-949-016-13261	Sequence 13261, A
C 36	217.8	13.9	46841	4	US-09-949-016-13466	Sequence 13466, A
C 37	217.6	13.9	246444	4	US-09-949-016-13113	Sequence 13113, A
38	217.6	13.9	321022	4	US-09-949-016-11852	Sequence 11852, A
39	217.6	13.9	321022	4	US-09-949-016-14166	Sequence 14166, A
C 40	217.2	13.8	101128	4	US-09-949-016-14293	Sequence 14293, A
C 41	216.8	13.8	10013	4	US-09-949-016-16474	Sequence 16474, A
C 42	216.6	13.8	44019	4	US-09-949-016-14902	Sequence 14902, A
C 43	216.4	13.8	461	3	US-09-404-879A-1	Sequence 1, Appli
C 44	216.4	13.8	461	3	US-09-404-879A-3	Sequence 3, Appli
C 45	216.4	13.8	461	4	US-09-338-933-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-13583
; Sequence 13583, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 13583

; LENGTH: 10818

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-13583

Query Match	99.8%	Score	1567.8	DB	4	Length	10818
Best Local Similarity	99.9%	Pred. No.	0				
Matches	1569	Conservative	0	Mismatches	2	Indels	0
Gaps	0						
QY	1	AGGGAGTGATGGGAATTGGAAGACAGGAACACAAATAGTCCAAAGCAATGGATTCTA	60				
Db	7243	AGGGAGTGATGGGAATTGGAAGACAGGAACACAAATAGTCCAAAGCAATGGATTCTA	7302				
QY	61	TTGGAGTGATTTGCCCCCTAGGAAGACACTGGCAATACCAAGGACATTTTGGTTGCTCA	120				
Db	7303	TTGGAGTGATTTGCCCCCTAGGAAGACACTGGCAATACCAAGGACATTTTGGTTGCTCA	7362				
QY	121	CACTATATGGAGGGGCACTTACTGGCACTAATGATAGATGCAAGTGTCTGTTCAAC	180				
Db	7363	CACTATATGGAGGGGCACTTACTGGCACTAATGATAGATGCAAGTGTCTGTTCAAC	7422				
QY	181	ATGCTATGATGCACAGCGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC	240				
Db	7423	ATGCTATGATGCACAGCGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC	7482				
QY	241	AGTCCAGATCGAGGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGGAG	300				
Db	7483	AGTCCAGATCGAGGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGGAG	7542				
QY	301	GTATAAGGATGGGTTGGTGAGAGATGGGAAGGAGTGTGTGTCCAGTAAGAAATAA	360				

Db	181	TCCCTCTCTCTCCCTTTTGGCTGTGTCCATCTCTGTCTTTTCTGCAITTTCTTCATCTC	122
Qy	962	TGTACTTTTCCATCTCTCTGTGTCTGTGTCCCATCTGTCTTCTCCATCTATGGCATCTCTGG	1021
Db	121	TGTACTTTTCCATCTCTCTGTGTGTGTGTCCCATCTGTCTTCTCCATCTATGGCATCTCTGG	62
Qy	1022	GTCTCTCATGTCTCTCTCTGCGCCACTTTGGCACACTCTGTGCTCTCTCATGCCCCCCTTT	1081
Db	61	GTCTCTCATGTCTCTCTCTGCGCCACTTTGGCACACTCTGTGCTCTCTCTGTGCCCCCCTTT	2
Qy	1082	C 1082	
Db	1	C 1	

```

RESULT 3
US-09-602-877A-94/c
; Sequence 94, Application US/09602877A
; Patent No. 8432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 94
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-94

```

Query Match	30.7%;	Score 482;	DB 3;	Length 735;
Best Local Similarity	100.0%;	Prod. No. 1.1e-119;		
Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1090	CAGGGTGATTCTGGGGGGCCCTGTGGTCTGCAAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	1149	
Db	500	CAGGGTGATTCTGGGGGGCCCTGTGGTCTGCAAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	441	
Qy	1150	GGGATTACCCCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGCAACCTCTGCAAGTTC	1209	
Db	440	GGGATTACCCCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGCAACCTCTGCAAGTTC	381	
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAAACTCCTGAGTCACTCCAGGACTCAGCAC	1269	
Db	380	ACCAAGTGGATCCAGGAAACCATCCAGGCCAAACTCCTGAGTCACTCCAGGACTCAGCAC	321	
Qy	1270	CCGGCATCCCACTGTGTGAGGACAGCCCTGACATCTCTTTTCAGACCCCTCATTCCTTC	1329	
Db	320	CCGGCATCCCACTGTGTGAGGACAGCCCTGACATCTCTTTTCAGACCCCTCATTCCTTC	261	
Qy	1330	CCAGAGATGTGAAAGTTTCATCTCTCCAGCCCGCTGACCCCATGTCTCTGGACTCAGG	1389	
Db	260	CCAGAGATGTGAAAGTTTCATCTCTCCAGCCCGCTGACCCCATGTCTCTGGACTCAGG	201	
Qy	1390	GTCTGCTTCCCCACATTTGGGCTGCACCGTGTCTCTAGTTTGAACCCCTGGGAAACAATTC	1449	
Db	200	GTCTGCTTCCCCACATTTGGGCTGCACCGTGTCTCTAGTTTGAACCCCTGGGAAACAATTC	141	
Qy	1450	CAAAACATGTCACAGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1509	
Db	140	CAAAACATGTCACAGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	81	
Qy	1510	TCAGGGCCCATTCCTTCTCTGCACTGTGACCCAAAATTTAGTCCAGAAATAAACTGAGA	1569	
Db	80	TCAGGGCCCATTCCTTCTCTGCACTGTGACCCAAAATTTAGTCCAGAAATAAACTGAGA	21	

Qy	1570 AG 1571	
Db	20 AG 19	

RESULT 4
 US-09-949-016-1841
 ; Sequence 1841, Application US/09949016
 ; Patent NO. 681239
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1841
 ; LENGTH: 1143
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-1841

Query Match	30.7%	Score 482	DB 4	Length 1143
Best Local Similarity	100.0%	Pred. No. 1.3e-119		
Matches 482	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1090	CAGGGTGATTTCTGGGGGCGCTTGGTCTGCAATGGCTCCCTGCAGGGACTCTGTGCTCTGG	1149	
Db	657	CAGGGTGATTTCTGGGGGCGCTTGGTCTGCAATGGCTCCCTGCAGGGACTCTGTGCTCTGG	716	
Qy	1150	GGAGATTACCTTGTGCCCCGCCCAACAGACGGGGTGTCTACACGAACCTCTGCAAGTTTC	1209	
Db	717	GGAGATTACCTTGTGCCCCGCCCAACAGACGGGGTGTCTACACGAACCTCTGCAAGTTTC	776	
Qy	1210	ACCAAGTGGATTCACAGGAACCAATCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCAC	1269	
Db	777	ACCAAGTGGATTCACAGGAACCAATCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCAC	836	
Qy	1270	CCGCGATCCCCACTGCTGCAGGGACAGCCCTGCACACTCCTTTACAGCCCTCATTCCTTTC	1329	
Db	837	CCGCGATCCCCACTGCTGCAGGGACAGCCCTGCACACTCCTTTACAGCCCTCATTCCTTTC	896	
Qy	1330	CCAGAGATGTTGAGAATGTTTCATCTCTCAGCCCCCTGACCCCCATGTCCTCTGGACTCAGG	1389	
Db	897	CCAGAGATGTTGAGAAATGTTTCATCTCTCAGCCCCCTGACCCCCATGTCCTCTGGACTCAGG	956	
Qy	1390	GTCTGCTTCCCCACATATGGGCTGACCGTGTCTCTAGTTGAAACCTTGGGAACAATTTTC	1449	
Db	957	GTCTGCTTCCCCACATATGGGCTGACCGTGTCTCTAGTTGAAACCTTGGGAACAATTTTC	1016	
Qy	1450	CAAAACTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGC	1509	
Db	1017	CAAAACTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGC	1076	
Qy	1510	TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569	
Db	1077	TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA	1136	
Qy	1570	AG 1571		
Db	1137	AG 1138		

RESULT 5
US-09-509-908-1

; Sequence 1, Application US/09509908
; Patent No. 6589770
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company, N/A N/A
; TITLE OF INVENTION: A Protease
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: T. David Reed
; STREET: 5299 Spring Grove Avenue
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45217-1087
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/509,908
; FILING DATE: 28-Feb-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, T David
; REGISTRATION NUMBER: 32,931
; REFERENCE/DOCKET NUMBER: AA-264F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-627-7025
; TELEFAX: 513-627-6333
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1499 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 291..1172
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 489..1172
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..290
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1173..1499
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-509-908-1

Query Match 30.7%; Score 482; DB 4; Length 1499;
Best Local Similarity 100.0%; Pred. No. 1.5e-119;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1149
DB 1014 CAGGGTATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1073
QY 1150 GGAGATTACCTTTGTGCCGGCCCAACAGACCGGGTGTCTACAGCAACCTTCGAAAGTTC 1209
DB 1074 GGAGATTACCTTTGTGCCGGCCCAACAGACCGGGTGTCTACAGCAACCTTCGAAAGTTC 1133
QY 1210 ACCAAGTGGATCCAGGAAACATCCAGGCCAATCTCTGAGTCAATCCAGGAGCTCAGACA 1269
DB 1134 ACCAAGTGGATCCAGGAAACATCCAGGCCAATCTCTGAGTCAATCCAGGAGCTCAGACA 1193
QY 1270 CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCAATTCCTTC 1329
DB 1194 CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCAATTCCTTC 1253
QY 1330 CCAGAGATGTGAGATGTTTCATCTCTCCAGCCCTGACCCCAATGTTCTCTGGACTCAGG 1389

DB 1254 CCAGAGATGTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTCGACTCAGG 1313
QY 1390 GTCTGCTTCCCCACATTCGGGCTGACCGTGTCTCTAGTTGAACCTCGGACAAATTC 1449
DB 1314 GTCCTTCCCCACATTCGGGCTGACCGTGTCTCTAGTTGAACCTCGGACAAATTC 1373
QY 1450 CAAAACCTGTCAGGGCGGGGGTTCCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
DB 1374 CAAAACCTGTCAGGGCGGGGGTTCCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1433
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569
DB 1434 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1493
QY 1570 AG 1571
DB 1494 AG 1495

RESULT 6

US-09-949-016-63415/c
; Sequence 63415, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63415
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63415

Query Match 30.6%; Score 480.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.3e-119;
Matches 480; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGAGGTATGGGAATTGAAGACAGGAAACACAAAATTAGTCCAAAGCGAATGGATTCTA 60
DB 481 AGGGAGGTATGGGAATTGAAGACAGGAAACACAAAATTAGTCCAAAGCGAATGGATTCTA 422
QY 61 TTGGGAGTGAATTCGCCCTTAGAAGACACTGGCAATACCAAGAGACATTTTGGTGTCA 120
DB 421 TTGGGAGTGAATTCGCCCTTAGAAGACACTGGCAATACCAAGAGACATTTTGGTGTCA 362
QY 121 CAACTATATGAGGGGCAATTAAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 180
DB 361 CAACTATATGAGGGGCAATTAAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 302
QY 181 ATGCTATGATGCACAGCGAGGCTCCACAAACAAACATTTATCCAGCTTCAGATGCCAC 240
DB 301 ATGCTATGATGCACAGCGAGGCTCCACAAACAAACATTTATCCAGCTTCAGATGCCAC 242
QY 241 AGTGCCAGATCGAGAAACCTCATCCAGGGCTGAGAACCGTATTTTTCAGAGGGAG 300
DB 241 AGTGCCAGATCGAGAAACCTCATCCAGGGCTGAGAACCGTATTTTTCAGAGGGAG 182
QY 301 GTATAGGATGGTGTGGTGGAGTGGGAAGGAGGTGTGTCTCCAGTAAGGAAATAA 360
DB 181 GTATAGGATGGTGTGGTGGAGTGGGAAGGAGGTGTGTCTCCAGTAAGGAAATAA 122
QY 361 GGCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGGGAGGAGATACACGATGA 420

Db 121 GGCCTGCACAGGCTGGAGGGAGAGTGCAGAGAGAAAGGAGCGGAGAGATACACAGTGA 62
QY 421 GGGAGACAGGCTGGAACAGAAAGTAGAGACGAAGATTGCGAGATGTGGAGAGGAAGGTCA 480
Db 61 GGGAGACAGGCTGGAACAGAAAGTAGAGACGAAGATTGCGAGATGTGGAGAGGAAGGTCA 2
QY 481 C 481
Db 1 C 1

RESULT 7
US-08-874-2
; Sequence 2, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: Filed Herewith
; APPLICATION NUMBER: US/08/824,874
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KERANOT02
; CLONE: 820694
US-08-824-874-2

Query Match 30.3%; Score 476; DB 2; Length 1476;
Best Local Similarity 100.0%; Pred. No. 6e-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGAATTCGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTCTCTGG 1149
Db 1001 CAGGGTGAATTCGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTCTCTGG 1060
QY 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209
Db 1061 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1120
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 1269
Db 1121 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 1180

QY 1270 CCGGCATCCCACTGCTGTCAGGGAGAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1329
Db 1181 CCGGCATCCCACTGCTGTCAGGGAGAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1240
QY 1330 CCAGAGATGTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1389
Db 1241 CCAGAGATGTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1300
QY 1390 GTCTGCTTCCCCACACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
Db 1301 GTCTGCTTCCCCACACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1360
QY 1450 CAAACTGTCCAGGGGGGGTTCGCTCAATCTCCCTGGGACATTTTCATCTCTCAAGC 1509
Db 1361 CAAACTGTCCAGGGGGGGTTCGCTCAATCTCCCTGGGACATTTTCATCTCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACT 1565
Db 1421 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACT 1476

RESULT 8
US-09-210-084-2
; Sequence 2, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,084
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KERANOT02
; CLONE: 820694
US-09-210-084-2

Query Match 30.3%; Score 476; DB 3; Length 1476;
Best Local Similarity 100.0%; Pred. No. 6e-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGAATTCGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTCTCTGG 1149

Db 1001 ||||| CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 1060
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCACTCTCGAAGTTC 1209
Db 1061 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCACTCTCGAAGTTC 1120
QY 1210 ACCAAGTGAATCCAGGAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1269
Db 1121 ACCAAGTGAATCCAGGAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1180
QY 1270 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1329
Db 1181 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGGCCCTGACCCCATGTCTCTGAGTCACTCCAGGACTCAGG 1389
Db 1241 CCAGAGATGTTGAGATGTTTCATCTCTCCAGGCCCTGACCCCATGTCTCTGAGTCACTCCAGGACTCAGG 1300
QY 1390 GTCTGCTTCCCACTGAGTGGGCTGACCGTGTCTCTAGTTGAACCTCGGGAACAATTC 1449
Db 1301 GTCTGCTTCCCACTGAGTGGGCTGACCGTGTCTCTAGTTGAACCTCGGGAACAATTC 1360
QY 1450 CAAAACCTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
Db 1361 CAAAACCTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAACT 1565
Db 1421 TCAGGGCCCATCCCTTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAACT 1476

RESULT 9

US-09-764-762-2
; Sequence 2, Application US/09764762
; Patent No. 6472195
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; LIBRARY: KERANOT02
; CLONE: 820694
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2

Query Match 30.3%; Score 476; DB 4; Length 1476;
Best Local Similarity 100.0%; Pred. No. 6e-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 1149
Db 1001 CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 1060
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCACTCTCGAAGTTC 1209
Db 1061 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCACTCTCGAAGTTC 1120
QY 1210 ACCAAGTGAATCCAGGAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1269
Db 1121 ACCAAGTGAATCCAGGAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1180
QY 1270 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1329
Db 1181 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGGCCCTGACCCCATGTCTCTGAGTCACTCCAGGACTCAGG 1389
Db 1241 CCAGAGATGTTGAGATGTTTCATCTCTCCAGGCCCTGACCCCATGTCTCTGAGTCACTCCAGGACTCAGG 1300
QY 1390 GTCTGCTTCCCACTGAGTGGGCTGACCGTGTCTCTAGTTGAACCTCGGGAACAATTC 1449
Db 1301 GTCTGCTTCCCACTGAGTGGGCTGACCGTGTCTCTAGTTGAACCTCGGGAACAATTC 1360
QY 1450 CAAAACCTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
Db 1361 CAAAACCTGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAACT 1565
Db 1421 TCAGGGCCCATCCCTTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAACT 1476

RESULT 10

US-09-280-116-1
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-1

Query Match 29.3%; Score 460; DB 3; Length 1504;
Best Local Similarity 99.6%; Pred. No. 1.2e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1090 CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCT-G 1148
Db 1015 CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 1074
QY 1149 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCACTCTCGAAGTTC 1208

Db 1075 GGAGATTACCTTTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTT 1134
Qy 1209 CACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTGTATCCAGGACTCAGCAC 1268
Db 1135 CACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTGTATCCAGGACTCAGCAC 1194
Qy 1269 ACCGGCATCCCACTCTGCGGAGACAGCCCTGACACTCTCTTTCAGACCTCATTCCTT 1328
Db 1195 ACCGGCATCCCACTCTGCGGAGACAGCCCTGACACTCTCTTTCAGACCTCATTCCTT 1254
Qy 1329 CCAGAGATGTGAGAAATGTTCACTCTCCAGGCCCTGACCCCAATGTCCTCTGAGACTCAG 1388
Db 1255 CCAGAGATGTGAGAAATGTTCACTCTCCAGGCCCTGACCCCAATGTCCTCTGAGACTCAG 1314
Qy 1389 GGTCTGTTCCCAACATGCGCTGACCGTGTCTCTAGTGTGAACCTGGGAACAATTT 1448
Db 1315 GGTCTGTTCCCAACATGCGCTGACCGTGTCTCTAGTGTGAACCTGGGAACAATTT 1374
Qy 1449 CCAAACTGTCCAGGGCGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAG 1508
Db 1375 CCAAACTGTCCAGGGCGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAG 1434
Qy 1509 CTCAGGCCCATCCCTTCTCTGCACTCTGACCCCAATTTAGT-CCAGAAATAAACTGA 1567
Db 1435 CTCAGGCCCATCCCTTCTCTGCACTCTGACCCCAATTTAGTCCCCAGAAATAAACTGA 1494
Qy 1568 GAAG 1571
Db 1495 GAAG 1498

RESULT 11

US-09-949-016-63414/c

; Sequence 63414, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

Db 73 ATGCTATGATGCACCGGAGGCTCCACAAACCAATATATCCAGCTTCAGATGCCAC 14
Qy 241 AGTCCCGAGATCG 253
Db 13 AGTCCCGAGATCG 1

RESULT 12

US-09-949-016-12799

; Sequence 12799, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-04-14

Query Match 14.2%; Score 223.4; DB 4; Length 68719;
Best Local Similarity 86.0%; Pred. No. 2.8e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Qy 570 TGAGGCCAGGATGTTGGCTCAGCCTGTAAATCCCAACTTT-GGAGGCTGAGGTGG 628
Db 57314 TGAGGCCAGGATGTTGGCTCAGCCTGTAAATCCCAACTTTGGGAGGCCAAGGCAGG 57373
Qy 629 CGAATCACTTTGAGGCCAGGAGTTTCGAGACCAAGCTGCGCAACATGTTGAACCCCTCTC 688
Db 57374 TGATCACTTGTAGGTGAGGAGTTCAAGACCAAGCTGCGCAACATGTTGAACCCCTCTC 57433
Qy 689 TACAAAAAATACAAAAATTTAGCCGGTGTGTGTATGGACACCTGTAGTCAAGCTAC 748
Db 57434 TACTAAAAATACAAAAATTTAGCCAGGATGTTGGTGGGCACTCTAATCCAGTGAC 57493
Qy 749 TTGGAGGCTGAGGCCAGGAGATTTGTTGAACCCGGGAGATGGAGGCTGAGGTGAGCTGA 808
Db 57494 TTGGAGGCTGAGGCCAGGAGATTTGTTGAACCCCTGGGAGACAGAGGTTGCAAGTGA 57553
Qy 809 GGTGAGGCCACTGGCTCCCAACTGGGCAACAGAGTAAAGTCCATCTCAAAAAA 868
Db 57554 GATCGTCCCACTGTACTTCCACCTGGGCAACAGAGGAGTCCCTCTCAGAAAAA 57613
Qy 869 A 869
Db 57614 A 57614

RESULT 13

US-09-949-016-14296

; Sequence 14296, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14296
; LENGTH: 68720
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14296

Query Match      14.2%; Score 223.4; DB 4; Length 68720;
Best Local Similarity 86.0%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 570 TGGAGGCCAGGCATGGTGGCTCAGCGCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGGG 628
Db 57314 TGGAGGCTGGGCACGGTGGCTCAAGCCTGTAATCCCAACACTTTTGGGAGGCCAAGGCAGG 57373

QY 629 CGAATCACTTGGCCAGGAGTTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCTCTC 688
Db 57374 TGGATCACCTGAGTCAAGGAGTTCAAGACCAAGCTGGCCAAACATGGCCAAACCCCTCTC 57433

QY 689 TACAAAAAATAACAAAAATTAGCCGGTGTGGTGTGATGACACCTGTAGTCACAGCTAC 748
Db 57434 TACTAAAAAATACAGAAATTTAGCCAGGCATGGTGTGGGACCTGTATCCAGTGAC 57493

QY 749 TTGGGAGGCTGAGCCAGGAGAATTTGCTTGAACCCGGGAGATGAGGCTGCAGTGAGCTGA 808
Db 57494 TTGGGAGGCTGAGCCAGGAGAATTTGCTTGAACCTGGGAGACAGAGTTGTCAGTGAGCTGA 57553

QY 809 GGTCAAGCCACTGGCTCCAACTGGCCACACAGATGAAGACTCCATCTCAAAAAA 868
Db 57554 GATCGTCCCACTGTACTCCACCTGGCCAAACAGAGAGGGAGTCCCTCTCAAAAAA 57613

QY 869 A 869
Db 57614 A 57614

RESULT 14
US-09-949-016-14575
; Sequence 14575, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14575
; LENGTH: 32278
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14575

Query Match      14.2%; Score 222.8; DB 4; Length 32278;
Best Local Similarity 85.8%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 569 TTGGAGGCCAGGCATGGTGGCTCAAGCCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGG 627
```

```
Db 13360 TTGTTGGCCAGGCGTAGTGGCTCACGCTGTATAATCCAGCACATTTGGGAGACCAAAAGTGG 13419
QY 628 GCGAATCACTTGGAGCCAGGAGTTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCTCT 687
Db 13420 GCAGATCACTTAAGGTTCAGGAGTTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCTCT 13479
QY 688 CTACAAAAAATAACAAAAATTAGCCGGTGTGGTGTGATGACACCTGTAGTCACAGCTA 747
Db 13480 CTACAAAAAATAACAAAAATTTAGCCAGGTGTGGTGGCAGGTGCCTGTAGTCCAGCTA 13539
QY 748 CTTGGGAGGCTGAGGCAGGAGAAATTCCTTGAACCCGGGAGATGGAGGCTGCAGTGAGCTG 807
Db 13540 CTTGGGAGGCTGAGGCAGGAGAAATTCCTTGAACCCGGCAGGTGGAGGTTTCAGTGAGCTG 13599
QY 808 AGTTCAGGCCACTCGCCTCCAACTGGCCAAACAGATGAAGACTCCATCTCAAAAAA 867
Db 13600 AGATGGCCACTGCACTCCAGCTGGGCAACAGAGGAGACTCCGCTCTCAAAACAACAC 13659
QY 868 AA 869
Db 13660 AA 13661

RESULT 15
US-09-128-155-17
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match      14.2%; Score 222.6; DB 3; Length 176373;
Best Local Similarity 84.0%; Pred. No. 7.2e-49;
Matches 263; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 558 AAAGTTGAACCTTGGAGCCAGGCATGGTGGCTCAGCGCTGTATATCCCAACACTTTT-GGA 616
Db 127012 AAAGAAAAAATATTATTGGCCAGGCACGGCGCTCATGCTATATATCCAGCACATTTGGGA 127071
QY 617 GGCTGAGGTGGGCGAATCACTTGAAGCCAGGAGTTTGAGACCAAGCTGGCCCAACATGGTG 676
Db 127072 GGCTGAGGAGGTGGATCACTCAGGTTCAGAGTTCGAGACCAAGCTGGCCCAACATGGTG 127131
QY 677 AAACCCCGTCTCTACAAAAAATAACAAAAATAGCCGGTGTGGTGTGATGGACACCTGT 736
Db 127132 AAACCCCATCTCTACTAAAAAATAACAAAAAATAGCCAGGTGTGGTGGGACCTGT 127191
QY 737 AGTCACAGCTACTTGGGAGGCTCAGGCAGGAGAAATTCCTTGAACCCGGGAGATGGAGGCT 796
Db 127192 AATCCAGCTACACGGGAGGCTCAGGCAGGAGAAATCGCTTGAACCCGGGAGGAGAGTT 127251
QY 797 GCAGTGAGGTGAGGTCAAGGCCACTGGCGCTCAACCTGGGCAACAGAGTAAAGACTCCATCT 856
```


Db 127252 GCAGTGAGCCCAAGTTCAGGCCACTGCACCCGCTGGGCAACAGAGACTTTGTCT 127311

Qy 857 CAAAAAAAAAAAA 869

Db 127312 CAAAAAAAAAAAA 127324

Search completed: July 3, 2005, 18:09:10
Job time : 181.302 secs

END PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 628.913 Seconds
(without alignments)

15667.668 Million cell updates/sec

Title: US-09-936-271C-13_COPY_10000_11570

Perfect score: 1571

Sequence: 1 agggaggtatgggaattga.....cccagaataaactgagag 1571

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

```

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	541	34.4	567	16	US-10-029-386-7604
C 2	482	30.7	586	9	US-09-954-531-849
C 3	482	30.7	586	9	US-09-954-531-1266
C 4	482	30.7	586	21	US-10-843-641A-1916
C 5	482	30.7	586	21	US-10-843-641A-2333
C 6	482	30.7	735	9	US-09-745-288-94
C 7	482	30.7	735	18	US-10-453-919-94

8	482	30.7	1260	17	US-10-172-118-1411	Sequence 1411, Ap
9	482	30.7	1260	17	US-10-295-027-515	Sequence 515, App
10	482	30.7	1260	17	US-10-173-999-53	Sequence 53, Appl
11	482	30.7	1260	18	US-10-342-887-1411	Sequence 1411, Ap
12	482	30.7	1381	9	US-09-739-907-52	Sequence 52, Appl
13	482	30.7	1381	11	US-09-938-671-52	Sequence 52, Appl
14	482	30.7	1381	11	US-10-935-098-52	Sequence 52, Appl
15	482	30.7	1400	21	US-10-956-157-7041	Sequence 7041, Ap
16	482	30.7	1439	9	US-09-739-907-53	Sequence 53, Appl
17	482	30.7	1439	11	US-09-938-671-53	Sequence 53, Appl
18	482	30.7	1439	21	US-10-935-098-53	Sequence 53, Appl
19	482	30.7	1481	20	US-10-723-860-5032	Sequence 5032, Ap
20	482	30.7	1516	9	US-09-739-907-37	Sequence 37, Appl
21	482	30.7	1516	11	US-09-938-671-37	Sequence 37, Appl
22	482	30.7	1516	21	US-10-935-098-37	Sequence 37, Appl
23	482	30.7	1536	21	US-10-956-157-1806	Sequence 1806, Ap
24	482	30.7	1570	9	US-09-989-722-308	Sequence 308, App
25	482	30.7	1570	9	US-09-989-723-308	Sequence 308, App
26	482	30.7	1570	9	US-09-989-279-308	Sequence 308, App
27	482	30.7	1570	9	US-09-989-727-308	Sequence 308, App
28	482	30.7	1570	9	US-09-989-731-308	Sequence 308, App
29	482	30.7	1570	9	US-09-989-732-308	Sequence 308, App
30	482	30.7	1570	9	US-09-991-073-308	Sequence 308, App
31	482	30.7	1570	9	US-09-990-442-308	Sequence 308, App
32	482	30.7	1570	9	US-09-991-163-308	Sequence 308, App
33	482	30.7	1570	9	US-09-993-604-308	Sequence 308, App
34	482	30.7	1570	9	US-09-990-456-308	Sequence 308, App
35	482	30.7	1570	9	US-09-989-721-308	Sequence 308, App
36	482	30.7	1570	9	US-09-992-598-308	Sequence 308, App
37	482	30.7	1570	9	US-09-989-293A-308	Sequence 308, App
38	482	30.7	1570	9	US-09-989-735-308	Sequence 308, App
39	482	30.7	1570	9	US-09-990-444-308	Sequence 308, App
40	482	30.7	1570	9	US-09-991-181-308	Sequence 308, App
41	482	30.7	1570	9	US-09-989-730-308	Sequence 308, App
42	482	30.7	1570	9	US-09-990-436-308	Sequence 308, App
43	482	30.7	1570	9	US-09-993-687-308	Sequence 308, App
44	482	30.7	1570	10	US-09-989-734-308	Sequence 308, App
45	482	30.7	1570	10	US-09-997-653-308	Sequence 308, App

ALIGNMENTS

RESULT 1

```

US-10-029-386-7604
; Sequence 7604, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7604
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
; OTHER INFORMATION: EST_HUMAN HIT: W73140.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: AF135028.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y337, EVALUE 6.00e-27

```

US-10-029-386-7604

Query Match 34.4%; Score 541; DB 16; Length 567;
Best Local Similarity 100.0%; Pred. No. 7.6e-150;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1031 GTCTCTTCTGCGCACATTTGCCACATCTGCGCTCTCTCATGCCCCCTTCTCTCTGC 1090
DB 1 GTCTCTTCTGCGCACATTTGCCACATCTGCGCTCTCTCATGCCCCCTTCTCTCTGC 60

QY 1091 AGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCCCTCCAGGACTGTGCTCTGG 1150
DB 61 AGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCCCTCCAGGACTGTGCTCTGG 120

QY 1151 GAGATTACCCCTTGTGCGCGGCCCAACAGACGCGGTGCTACAGAACTCTGCAAGTTCA 1210
DB 121 GAGATTACCCCTTGTGCGCGGCCCAACAGACGCGGTGCTACAGAACTCTGCAAGTTCA 180

QY 1211 CCAAGTGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 1270
DB 181 CCAAGTGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 240

QY 1271 CGGATCCCACTCTGCGAGGACAGCCCTGACATCTCTTCCAGACCCTCATCTCTCC 1330
DB 241 CGGATCCCACTCTGCGAGGACAGCCCTGACATCTCTTCCAGACCCTCATCTCTCC 300

QY 1331 CAGAGATGTTGAGAACTGTTCACTCTCCAGCCCTGACCCATGTTCTCTGGACTCAGGG 1390
DB 301 CAGAGATGTTGAGAACTGTTCACTCTCCAGCCCTGACCCATGTTCTCTGGACTCAGGG 360

QY 1391 TCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAAACCCCTGGGAAACAATTTCC 1450
DB 361 TCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAAACCCCTGGGAAACAATTTCC 420

QY 1451 AAAACTGTCAGGGCGGGGTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGCT 1510
DB 421 AAAACTGTCAGGGCGGGGTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGCT 480

QY 1511 CAGGGCCCATCCCTTCTCTCAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGAA 1570
DB 481 CAGGGCCCATCCCTTCTCTCAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGAA 540

QY 1571 G 1571
DB 541 G 541

RESULT 2

US-09-954-531-849/c
; Sequence 849, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Candi
; FILE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 849
; LENGTH: 586
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n=a,t,g or c
US-09-954-531-849

Query Match 30.7%; Score 482; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCTCGAGGACTCGTGTCTCTGG 1149
DB 501 CAGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCTCGAGGACTCGTGTCTCTGG 442

QY 1150 GAGATTACCCCTTGTGCGCGGCCCAACAGACGCGGTGCTACAGAACTCTGCAAGTTTC 1209
DB 441 GAGATTACCCCTTGTGCGCGGCCCAACAGACGCGGTGCTACAGAACTCTGCAAGTTTC 382

QY 1210 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 1269
DB 381 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 322

QY 1270 CCGGCATCCCCACTCTCTGAGGACAGCCCTGACATCTCTTTCAGACCCCTCAATTCCTTC 1329
DB 321 CCGGCATCCCCACTCTCTGAGGACAGCCCTGACATCTCTTTCAGACCCCTCAATTCCTTC 262

QY 1330 CCAGAGATGTTGAGAACTGTTCACTCTCCAGCCCTGACCCATGTTCTCTGGACTCAGG 1389
DB 261 CCAGAGATGTTGAGAACTGTTCACTCTCCAGCCCTGACCCATGTTCTCTGGACTCAGG 202

QY 1390 GTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAAACCCCTGGGAAACAATTTTC 1449
DB 201 GTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAAACCCCTGGGAAACAATTTTC 142

QY 1450 CAAAACCTGTCAGGGCGGGGTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 1509
DB 141 CAAAACCTGTCAGGGCGGGGTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 82

QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 1569
DB 81 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 22

QY 1570 AG 1571
DB 21 AG 20

RESULT 3

US-09-954-531-1266/c
; Sequence 1266, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Candi
; FILE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1266
; LENGTH: 586

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n=a,t,g or c
US-09-936-271c-13_copy_10000_11570.rnpb

Query Match      30.7%; Score 482; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGCTGATTTCTGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149
Db      |||
QY 501  CAGGCTGATTTCTGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 442
Db      |||
QY 1150 GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db      |||
QY 441  GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
Db      |||
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1269
Db      |||
QY 381  ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 322
Db      |||
QY 1270 CCGGCATCCCCACCTGCTGCAGGAGACCGCTGACCTCCCTTCAGACCCCTCATTCCTTC 1329
Db      |||
QY 321  CCGGCATCCCCACCTGCTGCAGGAGACCGCTGACCTCCCTTCAGACCCCTCATTCCTTC 262
Db      |||
QY 1330 CCAGAGATTGAGATGTTGATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
Db      |||
QY 261  CCAGAGATTGAGATGTTGATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 202
Db      |||
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
Db      |||
QY 201  GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 142
Db      |||
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 1509
Db      |||
QY 141  CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 82
Db      |||
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1569
Db      |||
QY 81   TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 22
Db      |||
QY 1570 AG 1571
Db      ||
      21 AG 20

RESULT 4
US-10-843-641A-1916/c
; Sequence 1916, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n=a,t,g or c
US-09-936-271c-13_copy_10000_11570.rnpb

Query Match      30.7%; Score 482; DB 21; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGCTGATTTCTGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149
Db      |||
QY 501  CAGGCTGATTTCTGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 442
Db      |||
QY 1150 GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db      |||
QY 441  GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
Db      |||
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1269
Db      |||
QY 381  ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 322
Db      |||
QY 1270 CCGGCATCCCCACCTGCTGCAGGAGACCGCTGACCTCCCTTCAGACCCCTCATTCCTTC 1329
Db      |||
QY 321  CCGGCATCCCCACCTGCTGCAGGAGACCGCTGACCTCCCTTCAGACCCCTCATTCCTTC 262
Db      |||
QY 1330 CCAGAGATTGAGATGTTGATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
Db      |||
QY 261  CCAGAGATTGAGATGTTGATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 202
Db      |||
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
Db      |||
QY 201  GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 142
Db      |||
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 1509
Db      |||
QY 141  CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 82
Db      |||
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1569
Db      |||
QY 81   TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 22
Db      |||
QY 1570 AG 1571
Db      ||
      21 AG 20

RESULT 5
US-10-843-641A-2333/c
; Sequence 2333, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
```

; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2333
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(586)
; OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-2333

Query Match 30.7%; Score 482; DB 21; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1149
DB 501 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 442

QY 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 1209
DB 441 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 382

QY 1210 ACCAAGTGGATCCAGGAACCAATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 1269
DB 381 ACCAAGTGGATCCAGGAACCAATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 322

QY 1270 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTTC 1329
DB 321 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTTC 262

QY 1330 CCAGAGATGTTGAGATGTTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGAGTCAAG 1389
DB 261 CCAGAGATGTTGAGATGTTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGAGTCAAG 202

QY 1390 GTCTGCTTCCCCACATTGGGCTGACCGTCTCTCTAGTTGAACCCCTGGGAACAATTTTC 1449
DB 201 GTCTGCTTCCCCACATTGGGCTGACCGTCTCTCTAGTTGAACCCCTGGGAACAATTTTC 142

QY 1450 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCTCCCTGGGCACTTTTCATCTCTCAAGC 1509
DB 141 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCTCCCTGGGCACTTTTCATCTCTCAAGC 82

QY 1510 TCAGGGCCCATCTCTTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTCTGAGA 1569
DB 81 TCAGGGCCCATCTCTTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTCTGAGA 22

QY 1570 AG 1571
DB 21 AG 20

RESULT 6
US-09-745-288-94/c
; Sequence 94, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 94
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-94

Query Match 30.7%; Score 482; DB 9; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.8e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1149
DB 500 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 441

QY 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 1209
DB 440 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 381

QY 1210 ACCAAGTGGATCCAGGAACCAATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 1269
DB 380 ACCAAGTGGATCCAGGAACCAATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 321

QY 1270 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTTC 1329
DB 320 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTTC 261

QY 1330 CCAGAGATGTTGAGATGTTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGAGTCAAG 1389
DB 260 CCAGAGATGTTGAGATGTTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGAGTCAAG 201

QY 1390 GTCTGCTTCCCCACATTGGGCTGACCGTCTCTCTAGTTGAACCCCTGGGAACAATTTTC 1449
DB 200 GTCTGCTTCCCCACATTGGGCTGACCGTCTCTCTAGTTGAACCCCTGGGAACAATTTTC 141

QY 1450 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCTCCCTGGGCACTTTTCATCTCTCAAGC 1509
DB 140 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCTCCCTGGGCACTTTTCATCTCTCAAGC 81

QY 1510 TCAGGGCCCATCTCTTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTCTGAGA 1569
DB 80 TCAGGGCCCATCTCTTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTCTGAGA 21

QY 1570 AG 1571
DB 20 AG 19

RESULT 7
US-10-453-919-94/c
; Sequence 94, Application US/10453919
; Publication No. US2004003230A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

;; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
;; FILE REFERENCE: 210121.446C7
;; CURRENT APPLICATION NUMBER: US/10/453,919
;; CURRENT FILING DATE: 2003-06-03
;; NUMBER OF SEQ ID NOS: 121
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 94
;; LENGTH: 735
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-10-453-919-94

Query Match 30.7%; Score 482; DB 18; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.8e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 1149
Db |||||
QY 500 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 441
Db |||||
QY 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db |||||
QY 440 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 381
Db |||||
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCTATCCAGGACTCAGACA 1269
Db |||||
QY 380 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCTATCCAGGACTCAGACA 321
Db |||||
QY 1270 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTCAGACCCCTCATTCCTTC 1329
Db |||||
QY 320 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTCAGACCCCTCATTCCTTC 261
Db |||||
QY 1330 CCAGAGATGTGAGATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389
Db |||||
QY 260 CCAGAGATGTGAGATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 201
Db |||||
QY 1390 GTCTGCTTCCCCACATTTGGCTGACCGTCTCTCTAGTGAACCCCTGGGAACAATTC 1449
Db |||||
QY 200 GTCTGCTTCCCCACATTTGGCTGACCGTCTCTCTAGTGAACCCCTGGGAACAATTC 141
Db |||||
QY 1450 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTCTATCTCTCAAGC 1509
Db |||||
QY 140 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTCTATCTCTCAAGC 81
Db |||||
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db |||||
QY 80 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 21
Db |||||
QY 1570 AG 1571
Db ||
QY 20 AG 19

RESULT 8
US-10-172-118-1411
;; Sequence 1411, Application US/10172118
;; Publication No. US20030224374A1
;; GENERAL INFORMATION:
;; APPLICANT: Dai, Hongyue
;; APPLICANT: He, Yudong
;; APPLICANT: Linsley, Peter
;; APPLICANT: Mao, Mao
;; APPLICANT: Roberts, Chris
;; APPLICANT: Van 't Veer, Laura
;; APPLICANT: Van de Vijver, Marc
;; APPLICANT: Bernards, Rene
;; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;; FILE REFERENCE: 9301-175-999
;; CURRENT APPLICATION NUMBER: US/10/172,118
;; CURRENT FILING DATE: 2002-06-14
;; PRIOR APPLICATION NUMBER: 60/380,770
;; PRIOR FILING DATE: 2002-05-14
;; NUMBER OF SEQ ID NOS: 2699

;; SEQ ID NO 1411
;; LENGTH: 1260
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: NM 012427
;; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1411

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 1149
Db |||||
QY 766 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 825
Db |||||
QY 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db |||||
QY 826 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 885
Db |||||
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCTATCCAGGACTCAGACA 1269
Db |||||
QY 886 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCTATCCAGGACTCAGACA 945
Db |||||
QY 1270 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTCAGACCCCTCATTCCTTC 1329
Db |||||
QY 946 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTCAGACCCCTCATTCCTTC 1005
Db |||||
QY 1330 CCAGAGATGTGAGATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389
Db |||||
QY 1006 CCAGAGATGTGAGATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1065
Db |||||
QY 1390 GTCTGCTTCCCCACATTTGGCTGACCGTCTCTCTAGTGAACCCCTGGGAACAATTC 1449
Db |||||
QY 200 GTCTGCTTCCCCACATTTGGCTGACCGTCTCTCTAGTGAACCCCTGGGAACAATTC 1125
Db |||||
QY 1450 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTCTATCTCTCAAGC 1509
Db |||||
QY 1126 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTCTATCTCTCAAGC 1185
Db |||||
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db |||||
QY 1186 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1245
Db |||||
QY 1570 AG 1571
Db ||
QY 1246 AG 1247

RESULT 9
US-10-295-027-515
;; Sequence 515, Application US/10295027
;; Publication No. US20030232350A1
;; GENERAL INFORMATION:
;; APPLICANT: Afar, Daniel
;; APPLICANT: Aziz, Natasha
;; APPLICANT: Ginsberg, Wendy M.
;; APPLICANT: Gish, Kurt C.
;; APPLICANT: Glynn, Richard
;; APPLICANT: Hevezi, Peter A.
;; APPLICANT: Mack, David H.
;; APPLICANT: Murray, Richard
;; APPLICANT: Watson, Susan R.
;; APPLICANT: Eos Biotechnology, Inc.
;; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
;; FILE REFERENCE: 018501-012500US
;; CURRENT APPLICATION NUMBER: US/10/295,027
;; CURRENT FILING DATE: 2002-11-13
;; PRIOR APPLICATION NUMBER: US 09/663,733
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 515
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-295-027-515

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGATTCGCGGGGCGCTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCTCTGG 1149
DB 766 CAGGGTGATTCGCGGGGCGCTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCTCTGG 825
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 1209
DB 826 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 885
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1269
DB 886 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 945
QY 1270 CCGGCATCCCACTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCCCTCAATTCCTTC 1329
DB 946 CCGGCATCCCACTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCCCTCAATTCCTTC 1005
QY 1330 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389
DB 1006 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1065
QY 1390 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC 1449
DB 1066 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC 1125
QY 1450 CAAAACCTGTCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
DB 1126 CAAAACCTGTCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1185
QY 1510 TCAGGGCCCATCTCTCTGACGCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTCTGAGA 1569
DB 1186 TCAGGGCCCATCTCTCTGACGCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTCTGAGA 1245
QY 1570 AG 1571
DB 1246 AG 1247

RESULT 10
US-10-173-999-53
; Sequence 53, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.

; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-173-999-53

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGATTCGCGGGGCGCTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCTCTGG 1149
DB 766 CAGGGTGATTCGCGGGGCGCTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCTCTGG 825
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 1209
DB 826 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 885
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1269
DB 886 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 945
QY 1270 CCGGCATCCCACTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCCCTCAATTCCTTC 1329
DB 946 CCGGCATCCCACTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCCCTCAATTCCTTC 1005
QY 1330 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389
DB 1006 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1065
QY 1390 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC 1449
DB 1066 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC 1125
QY 1450 CAAAACCTGTCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
DB 1126 CAAAACCTGTCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1185
QY 1510 TCAGGGCCCATCTCTCTGACGCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTCTGAGA 1569
DB 1186 TCAGGGCCCATCTCTCTGACGCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTCTGAGA 1245
QY 1570 AG 1571
DB 1246 AG 1247

RESULT 11
US-10-342-887-1411
; Sequence 1411, Application US/10342887
; Publication No. US200400058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.

APPLICANT: Mao, Mao
APPLICANT: Roberts, Christopher J.
APPLICANT: Van 't Veer, Laura Johanna
APPLICANT: Van de Vijver, Marc J.
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-188-999
CURRENT APPLICATION NUMBER: US/10/342,887
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/298,918
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/380,710
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: 10/172,118
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1411
LENGTH: 1260
TYPE: DNA
ORGANISM: Homo sapiens
US-10-342-887-1411

Query Match 30.7%; Score 482; DB 18; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1090	CAGGGTGATTC	CGGGGGCCCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG	1149
Db	766	CAGGGTGATTC	CGGGGGCCCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG	825
Qy	1150	GGAGATTACCT	TGTGCGCGGCCCAACAGACCGGGTGTCTACAGACCTCTGCAAGTTC	1209
Db	826	GGAGATTACCT	TGTGCGCGGCCCAACAGACCGGGTGTCTACAGACCTCTGCAAGTTC	885
Qy	1210	ACCAAGTGGAT	CCAGGAAACCATCCAGGCAACTCTCTAGTTCATCCAGGACTCAGCACA	1269
Db	886	ACCAAGTGGAT	CCAGGAAACCATCCAGGCAACTCTCTAGTTCATCCAGGACTCAGCACA	945
Qy	1270	CCGGCATCCCC	ACCTGCTCAGGAGACGCCCTGACATCTCTTTTCAGACCTCTCATCTCTTC	1329
Db	946	CCGGCATCCCC	ACCTGCTCAGGAGACGCCCTGACATCTCTTTTCAGACCTCTCATCTCTTC	1005
Qy	1330	CCAGAGATGTT	GAGAAATGTTTCACTCTCCAGGCCCTGACCCCATGTCTCTGACTCAGG	1389
Db	1006	CCAGAGATGTT	GAGAAATGTTTCACTCTCCAGGCCCTGACCCCATGTCTCTGACTCAGG	1065
Qy	1390	GTCTGCTTCCC	CCACATTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC	1449
Db	1066	GTCTGCTTCCC	CCACATTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC	1125
Qy	1450	CAAACTGTCC	AGGCGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC	1509
Db	1126	CAAACTGTCC	AGGCGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC	1185
Qy	1510	TCAGGGCCAT	CCCTTCTCTGACGCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569
Db	1186	TCAGGGCCAT	CCCTTCTCTGACGCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1245
Qy	1570	AG	1571	
Db	1246	AG	1247	

RESULT 12
US-09-936-671-52
Sequence 52, Application US/09739907
Patent No. US20010012889A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/739,907
CURRENT FILING DATE: 2000-12-20

PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
PRIOR FILING DATE: 1998-01-07
NUMBER OF SEQ ID NOS: 196
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 52
LENGTH: 1381
TYPE: DNA
ORGANISM: Homo sapiens
US-09-739-907-52

Query Match 30.7%; Score 482; DB 9; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1090	CAGGGTGATTC	CGGGGGCCCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG	1149
Db	852	CAGGGTGATTC	CGGGGGCCCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG	911
Qy	1150	GGAGATTACCT	TGTGCGCGGCCCAACAGACCGGGTGTCTACAGACCTCTGCAAGTTC	1209
Db	912	GGAGATTACCT	TGTGCGCGGCCCAACAGACCGGGTGTCTACAGACCTCTGCAAGTTC	971
Qy	1210	ACCAAGTGGAT	CCAGGAAACCATCCAGGCAACTCTCTAGTTCATCCAGGACTCAGCACA	1269
Db	972	ACCAAGTGGAT	CCAGGAAACCATCCAGGCAACTCTCTAGTTCATCCAGGACTCAGCACA	1031
Qy	1270	CCGGCATCCCC	ACCTGCTCAGGAGACGCCCTGACATCTCTTTTCAGACCTCTCATCTCTTC	1329
Db	1032	CCGGCATCCCC	ACCTGCTCAGGAGACGCCCTGACATCTCTTTTCAGACCTCTCATCTCTTC	1091
Qy	1330	CCAGAGATGTT	GAGAAATGTTTCACTCTCCAGGCCCTGACCCCATGTCTCTGACTCAGG	1389
Db	1092	CCAGAGATGTT	GAGAAATGTTTCACTCTCCAGGCCCTGACCCCATGTCTCTGACTCAGG	1151
Qy	1390	GTCTGCTTCCC	CCACATTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC	1449
Db	1152	GTCTGCTTCCC	CCACATTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTC	1211
Qy	1450	CAAACTGTCC	AGGCGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC	1509
Db	1212	CAAACTGTCC	AGGCGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC	1271
Qy	1510	TCAGGGCCAT	CCCTTCTCTGACGCTTTCAGCCTTTCAGCCAAATTTAGTCCAGAAATAAATCTGAGA	1569
Db	1272	TCAGGGCCAT	CCCTTCTCTGACGCTTTCAGCCTTTCAGCCTTTCAGCCAAATTTAGTCCAGAAATAAATCTGAGA	1331
Qy	1570	AG	1571	
Db	1332	AG	1333	

RESULT 13
US-09-938-671-52
Sequence 52, Application US/09938671
Publication No. US2004002066A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/938,671
CURRENT FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07

```

; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-671-52

Query Match      30.7%; Score 482; DB 11; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1149
Db 852 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 911

QY 1150 GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 912 GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 971

QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCCTGAGTCTATCCAGGACTCAGCACA 1269
Db 972 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCCTGAGTCTATCCAGGACTCAGCACA 1031

QY 1270 CCGGCATCCCCACCTGCTGCAGGAGACAGCCCTGACACTCCCTTTTCAGACCCCTCATTCCTTC 1329
Db 1032 CCGGCATCCCCACCTGCTGCAGGAGACAGCCCTGACACTCCCTTTTCAGACCCCTCATTCCTTC 1091

QY 1330 CCAGAGATGTTGAGATGTTTCACTCTCCAGGCCCTGACCCCAATGCTCTCGACTCAGG 1389
Db 1092 CCAGAGATGTTGAGATGTTTCACTCTCCAGGCCCTGACCCCAATGCTCTCGACTCAGG 1151

QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1449
Db 1152 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1211

QY 1450 CAAAACCTGTCAGGCGGGGGTGGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
Db 1212 CAAAACCTGTCAGGCGGGGGTGGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1271

QY 1510 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1569
Db 1272 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1331

QY 1570 AG 1571
Db 1332 AG 1333
```

```

RESULT 14
US-10-935-098-52
; Sequence 52, Application US/10935098
; Publication No. US20050042667A1
; GENERAL INFORMATION:
; APPLICANT: Lafleur et al.
; FILE REFERENCE: P2022P1C3
; CURRENT APPLICATION NUMBER: US/10/935,098
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: 09/938,671
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/739,907
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: PCT/US99/00108
; PRIOR FILING DATE: 1999-01-06
;
```

```

; PRIOR APPLICATION NUMBER: 60/070,657
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-935-098-52

Query Match      30.7%; Score 482; DB 21; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1149
Db 852 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 911

QY 1150 GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 912 GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 971

QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCCTGAGTCTATCCAGGACTCAGCACA 1269
Db 972 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCCTGAGTCTATCCAGGACTCAGCACA 1031

QY 1270 CCGGCATCCCCACCTGCTGCAGGAGACAGCCCTGACACTCCCTTTTCAGACCCCTCATTCCTTC 1329
Db 1032 CCGGCATCCCCACCTGCTGCAGGAGACAGCCCTGACACTCCCTTTTCAGACCCCTCATTCCTTC 1091

QY 1330 CCAGAGATGTTGAGATGTTTCACTCTCCAGGCCCTGACCCCAATGCTCTCGACTCAGG 1389
Db 1092 CCAGAGATGTTGAGATGTTTCACTCTCCAGGCCCTGACCCCAATGCTCTCGACTCAGG 1151

QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1449
Db 1152 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1211

QY 1450 CAAAACCTGTCAGGCGGGGGTGGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
Db 1212 CAAAACCTGTCAGGCGGGGGTGGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1271

QY 1510 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1569
Db 1272 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1331

QY 1570 AG 1571
Db 1332 AG 1333

RESULT 15
US-10-956-157-7041
; Sequence 7041, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7041
; LENGTH: 1400
;
```

TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-7041

Query Match 30.7%; Score 482; DB 21; Length 1400;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1090	CAGGGTGATTTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCTGG	1149
Db	881	CAGGGTGATTTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCTGG	940
Qy	1150	GGAGATTACCTTTGTGCCCCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC	1209
Db	941	GGAGATTACCTTTGTGCCCCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC	1000
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA	1269
Db	1001	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA	1060
Qy	1270	CCGGCATCCCACTGCTGAGGAGACGCCCTGACACTCCTTTGAGACCCCTCATTCCTTC	1329
Db	1061	CCGGCATCCCACTGCTGAGGAGACGCCCTGACACTCCTTTGAGACCCCTCATTCCTTC	1120
Qy	1330	CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1389
Db	1121	CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1180
Qy	1390	GTCTGCTTCCCCCACAATTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC	1449
Db	1181	GTCTGCTTCCCCCACAATTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC	1240
Qy	1450	CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC	1509
Db	1241	CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC	1300
Qy	1510	TCAGGGCCCATCCCTTCTCTGCACTCTGACCCAAATTTAGTCCAGAAATAAATGAGA	1569
Db	1301	TCAGGGCCCATCCCTTCTCTGCACTCTGACCCAAATTTAGTCCAGAAATAAATGAGA	1360
Qy	1570	AG 1571	
Db	1361	AG 1362	

Search completed: July 4, 2005, 02:13:09
Job time : 630.913 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2005, 00:19:06 ; Search time 3415.96 Seconds
(without alignments)
17505.766 Million cell updates/sec

Title: US-09-936-271c-13_COPY_10000_11570

Perfect score: 1571

Sequence: 1 agggagtgatgggaattga.....cccgaaataaactgagaag 1571

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	482	30.7	539	2	BF513278 UI-H-BW1-
C 2	482	30.7	586	7	W73140 zd55e11.s1
C 3	475.8	30.3	809	4	BG680075 602628224
C 4	455	29.0	478	1	AA862032 0146e09.s
C 5	443.8	28.2	802	4	BG682309 602629626
C 6	440	28.0	953	2	BG745465 601579834
C 7	431.8	27.5	666	4	BM842155 K-EST0119
C 8	417	26.5	645	1	AI002163 0t42h05.s
C 9	383.6	24.4	677	2	BE898385 601681219
C 10	381.8	24.3	453	7	N80762 za98f06.s1
C 11	355.8	22.6	363	2	AW105502 xd53g06.x
C 12	348	22.2	590	4	BM840511 K-EST0117
C 13	333.4	21.2	577	4	BM838406 K-EST0114
C 14	322.8	20.5	430	7	W68361 zd34h08.s1
C 15	317	20.2	388	7	W68496 zd34h08.r1
C 16	314	20.0	334	2	BF514439 UI-H-BW1-
C 17	291	18.5	533	4	BM840509 K-EST0117
C 18	278.8	17.7	523	4	BM841293 K-EST0118
C 19	277.2	17.6	320	2	AW380655 RC4-H7027
C 20	276.6	17.6	411	2	AW801647 ILS-UM006
C 21	273	17.4	512	4	BM830263 K-EST0103
C 22	270.4	17.2	526	4	BM841697 K-EST0118
C 23	245	15.6	498	7	W73168 zd55e11.r1
C 24	229.8	14.6	420	6	CB298043 220013_re

ALIGNMENTS

RESULT 1
BF513278/c

LOCUS

DEFINITION BF513278 539 bp mRNA linear EST 07-DEC-2000
UI-H-BW1-amo-e-03-0-UI.s1 NCI CGAP_Sub7 Homo sapiens cDNA clone
IMAGE:3070564 3', mRNA sequence.

ACCESSION BF513278

VERSION BF513278.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 539)

AUTHORS NCI-CGAP

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:

NCI-CGAP clone distribution information can be found through the

I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. 539

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3070564"

/lab_host="NCI CGAP Sub7"

/notes="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub7

is a subtracted library derived from NCI CGAP Sub6. The

single-stranded DNA preparation of NCI CGAP Sub6 was used

as a tracer in a subtractive hybridization with a driver

comprising: the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM

3334-3337, 3682-3683, 3798-3803 IMAGE Clones

1323276-1323911, 1456008-1456775, 1500552-1502855);

NCI CGAP Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778

(IMAGE Clones 1323912-1325831, 1471368-1472903,

1492104-1493255); NCI CGAP Lu5 pool 1 LLAM 3575-3582,
3851-3854 (IMAGE Clonoids 1414920-1417991,
1520904-1522439); NCI CGAP GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 - (IMAGE Clonoids
1257096-1258631, 1469064-1470983, 1475592-1476743);
NCI CGAP Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068
(IMAGE Clonoids 985608-986759, 1101192-1101959,
1217928-1220615); NCI CGAP Co10 pool 1 LLAM 2644-2653,
2871-2872 (IMAGE Clonoids
1057416-1061255, 1144584-1145351). (6% of the driver
population), plus a pool of 3,840 arrayed clones from
NCI CGAP Sub1 (IMAGE Clonoids 2708616-2710535) and
NCI CGAP Sub2 (IMAGE Clonoids 2710536-2712455) (4% of
the driver population), plus a pool of 11,136 clones from
NCI CGAP Sub3 (IMAGE Clonoids 2712456-2723591) (10% of
the driver population), plus a pool of 5,472 clones from
NCI CGAP Sub4 (IMAGE Clonoids 2723592-2729326) (40% of the
driver population), plus a pool of 4032 clones from
NCI CGAP Sub6 (IMAGE Clonoids 2728969-2733190) (40% of the
driver population). Subtraction was performed as
previously described (Bonaldo, Lennon & Soares (1996):
Normalization and Subtraction: Two Approaches To
Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG TISSUE=lung
TAG LIB=NCI CGAP-Lu13
TAG SEQ=GCCGG"

ORIGIN

Query Match 30.7%; Score 482; DB 2; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.6e-65;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1090	CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG	1149
Db	508	CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG	449
QY	1150	GGAGATTACCTTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTTCCAAATTC	1209
Db	448	GGAGATTACCTTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTTCCAAATTC	389
QY	1210	ACCAAGTGGATCCAGGAACATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGACA	1269
Db	388	ACCAAGTGGATCCAGGAACATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGACA	329
QY	1270	CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC	1329
Db	328	CCGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC	269
QY	1330	CCAGAGATGTGAGATGTTTCTCTTCAGCCCTGACCCCAATGTTCTCTGAGCTCAGG	1389
Db	268	CCAGAGATGTGAGATGTTTCTCTTCAGCCCTGACCCCAATGTTCTCTGAGCTCAGG	209
QY	1390	GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGGACATTTTC	1449
Db	208	GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGGACATTTTC	149
QY	1450	CAAAACTGTCCAGGGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTCATCTCTCAAGC	1509
Db	148	CAAAACTGTCCAGGGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTCATCTCTCAAGC	89
QY	1510	TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCCAAGAAATAACTGAGA	1569
Db	88	TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCCAAGAAATAACTGAGA	29
QY	1570	AG 1571	
Db	28	AG 27	
RESULT 2			
W73140/c			
LOCUS	W73140	586 bp	linear EST 16-OCT-1996
DEFINITION	zds5e11.s1 Soares_fetal_heart_NbH19W	Homo sapiens	cdna clone

IMAGE:344588 3' similar to PIR:A53968 A53968 serine proteinase SCCE
precursor - human ;, mRNA sequence.
W73140
VERSION W73140.1 GI:1383275
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 586)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 761 Std Error: 0.00
Seq primer: ETPRIMER
High quality sequence stop: 428.
Location/Qualifiers
1..586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1269963"
/db_xref="taxon:9606"
/clone="IMAGE:344588"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NbH19W"
/notes="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTCAAGTGGAGCGCCGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."

ORIGIN

Query Match 30.7%; Score 482; DB 7; Length 586;
Best Local Similarity 100.0%; Pred. No. 4.6e-65;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1090	CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG	1149
Db	501	CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG	442
QY	1150	GGAGATTACCTTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTTCTCAAGTTC	1209
Db	441	GGAGATTACCTTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTTCTCAAGTTC	382
QY	1210	ACCAAGTGGATCCAGGAACATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGACA	1269
Db	381	ACCAAGTGGATCCAGGAACATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGACA	322
QY	1270	CGGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC	1329
Db	321	CGGGCATCCCACTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCTCTATCTCTTC	262

```
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
Db 261 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 202
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
Db 201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 142
QY 1450 CAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 1509
Db 141 CAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 82
QY 1510 TCAGGGCCCATCCCTTCTCTCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACTCAGA 1569
Db 81 TCAGGGCCCATCCCTTCTCTCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACTCAGA 22
QY 1570 AG 1571
Db 21 AG 20

RESULT 3
LOCUS BG680075 809 bp mRNA linear EST 01-MAY-2001
DEFINITION 602628224F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753372 5',
mRNA sequence.
ACCESSION BG680075
VERSION BG680075.1 GI:13911472
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 809)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10612 row: o column: 05
High quality sequence stop: 707.
FEATURES
source
1..809
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753372"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCWV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5Kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 30.3%; Score 475.8; DB 4; Length 809;
Best Local Similarity 99.6%; Pred. No. 3.8e-64;
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1084 CTCCTGAGGTGATCTTGGGGCGCTGTGTCTGCAATGCTCCCTGCGGGACTCGTG 1143
Db 219 CTCCTGAGGTGATCTTGGGGCGCTGTGTCTGCAATGCTCCCTGCGGGACTCGTG 278
QY 1144 TCCTGGGAGATTACCTTGTGCGCGGCCCAAGACCGGGTGTCTACAGAACTCTGC 1203
```

```
Db 279 TCCTGGGAGATTACCTTGTGCGCGGCCCAAGACCGGGTGTCTACAGAACTCTGC 338
QY 1204 AAGTTCCACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCAATCCCAAGACTC 1263
Db 339 AAGTTCCACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCAATCCCAAGACTC 398
QY 1264 AGCACACCGGCATCCCACTGCTGACGGGACAGCCCTGACACTCTCTTTCAGACCCCTCAT 1323
Db 399 AGCACACCGGCATCCCACTGCTGACGGGACAGCCCTGACACTCTCTTTCAGACCCCTCAT 458
QY 1324 TCCTTCCACAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGA 1383
Db 459 TCCTTCCACAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGA 518
QY 1384 CTAGGGTCTGTCTCCCACTGCTGAGCGTGTCTCTCTAGTTGAACCCCTGGGAAC 1443
Db 519 CTCAGGGTCTGTCTCCCACTGCTGAGCGTGTCTCTCTAGTTGAACCCCTGGGAAC 578
QY 1444 AATTTCCAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCC 1503
Db 579 AATTTCCAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCC 638
QY 1504 TCAAGCTCAGGGCCCATCCCTTCTCTCAGCTCTGACCCCAAAATTTAGTCCAGAAATAA 1562
Db 639 TCAAGCTCAGGGCCCATCCCTTCTCTCAGCTCTGACCCCAAAATTTAGTCCAGAAATAA 697

RESULT 4
LOCUS AA862032 478 bp mRNA linear EST 13-APR-1999
DEFINITION O146809 s1 NCI_CGAP_HN3 Homo sapiens cDNA clone IMAGE:1485736 3',
similar to SW:SCCE_HUMAN P49862 STRATUM CORNEUM CHYMOTRIPTIC ENZYME
PRECUSOR ; mRNA sequence.
ACCESSION AA862032
VERSION AA862032.1 GI:2954511
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 478)
AUTHORS NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: John Enaley, M.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1647 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 233.
FEATURES
source
1..478
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1485736"
/tissue_type="squamous cell carcinoma from base of tongue"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_HN3"
/notes="Organ: tongue; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.0 kb. 5' adaptor sequence:
5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
(GA)10ACTAGTCTCGAGTTTTTTTTTTTTTTT 3'"
ORIGIN
```

Query Match 29.0%; Score 455; DB 1; Length 478;
Best Local Similarity 99.8%; Pred. No. 7.5e-61;
Matches 466; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1105 GGGCTGTGTCTGCAATGCTCTCTGAGGACTGTGTCCTGGGGAGATTACCTTTGT 1164
Db 478 GGGCTGTGTCTGCAATGCTCTCTGAGGACTGTGTCCTGGGGAGATTACCTTTGT 420
QY 1165 GCCGGGCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTTCCACCAAGTGGATCCAG 1224
Db 419 GCCGGGCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTTCCACCAAGTGGATCCAG 360
QY 1225 GAAACCATCCAGGCCAACTCTCTGAGTCATCCAGGACTCAGACACACCGGATCCCCACCT 1284
Db 359 GAAACCATCCAGGCCAACTCTCTGAGTCATCCAGGACTCAGACACACCGGATCCCCACCT 300
QY 1285 GCTGCGAGGACAGCCCTGACACTCTCTTTTCAGACCTCATTTCTTCCAGAGATGTTGAGA 1344
Db 299 GCTGCGAGGACAGCCCTGACACTCTTTTCAGACCTCATTTCTTCCAGAGATGTTGAGA 240
QY 1345 ATGTTTATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGGGTCTGTTCCCCAC 1404
Db 239 ATGTTTATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGGGTCTGTTCCCCAC 180
QY 1405 ATTGGGCTGACCGGTCTCTCTAGTTGAACCCCTGGGAACAATTTCCAAAACCTGTCCAGGG 1464
Db 179 ATTGGGCTGACCGGTCTCTCTAGTTGAACCCCTGGGAACAATTTCCAAAACCTGTCCAGGG 120
QY 1465 CGGGGGTGGCTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGCTCAGGGCCCATCCCT 1524
Db 119 CGGGGGTGGCTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGCTCAGGGCCCATCCCT 60
QY 1525 TCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGAG 1571
Db 59 TCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGAG 13

RESULT 5
BG682309 802 bp mRNA linear EST 01-MAY-2001
LOCUS 602629626F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754263 5',
DEFINITION mRNA sequence.
ACCESSION BG682309
VERSION BG682309.1 GI:13913706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 802)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10615 row: d column: 08
High quality sequence stop: 802.
Location/Qualifiers
1. 802
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sali; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
Query Match 28.2%; Score 443.8; DB 4; Length 802;
Best Local Similarity 99.3%; Pred. No. 3.5e-59;
Matches 456; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1090 CAGGGTGATTCGGGGGGCTGTGTGTCGCAATGGCTCCCTGAGGAGCTCGTGTCTCTGG 1149
Db 333 CAGGGTGATTCGGGGGGCTGTGTGTCGCAATGGCTCCCTGAGGAGCTCGTGTCTCTGG 392
QY 1150 GGAGATTACCTTTGTCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209
Db 393 GGAGATTACCTTTGTCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 452
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCAACCTCTCTGAGTCTATCCAGGACTCAGCACA 1269
Db 453 ACCAAGTGGATCCAGGAAACCATCCAGGCAACCTCTCTGAGTCTATCCAGGACTCAGCACA 512
QY 1270 CCGGCATCCCACTCTCTGAGGAGAGCCCTGACACTCTTTTCAGACCTCTCATTCCTTC 1329
Db 513 CCGGCATCCCACTCTCTGAGGAGAGCCCTGACACTCTCTTTTCAGACCTCTCATTCCTTC 572
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 1389
Db 573 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 632
QY 1390 GTCTGCTTCCCCCACA-TTGGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTT 1448
Db 633 GTCTGCTTCCCCCACA-TTGGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTT 692
QY 1449 CCAAAACTGTCCAGGGCGGGGTGGCTCTCAATCTCCCTGGGGCACTTTTCATCTCTCAAG 1508
Db 693 CCAAAACTGTCCAGGGCGGGGTGGCTCTCAATCTCCCTGGGGCACTTTTCATCTCTCAAG 752
QY 1509 CTCAGGGCCATCCCTCTCTCTGAGCTCTGACCCCAATTT 1547
Db 753 CTCAGGGCCATCCCTCTCTCTGAGCTCTGACCCCAATTT 791

RESULT 6
BE745465 953 bp mRNA linear EST 15-SEP-2000
LOCUS 601579834F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3928640 5',
DEFINITION mRNA sequence.
ACCESSION BE745465
VERSION BE745465.1 GI:10159457
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 953)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW760 row: c column: 09
High quality sequence stop: 685.
Location/Qualifiers
1. 953
/organism="Homo sapiens"
/mol_type="mRNA"


```
/db_xref="taxon:9606"  
/clone="IMAGE:3928640"  
/tissue_type="adenocarcinoma cell line"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NIH_MGC_9"  
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCAGGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```
Query Match 28.0%; Score 440; DB 2; Length 953;  
Best Local Similarity 99.6%; Pred. No. 1.3e-58;  
Matches 462; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
  
QY 1090 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCGCAGGAGACTCGTCTCTGG 1149  
Db 229 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCGCAGGAGACTCGTCTCTGG 289  
  
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 1209  
Db 289 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 348  
  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTTCATCCAGGACTCAGCAC 1269  
Db 349 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTTCATCCAGGACTCAGCAC 408  
  
QY 1270 CCGGCATCCCACTGCTGAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329  
Db 409 CCGGCATCCCACTGCTGAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 468  
  
QY 1330 CCAGAGATTGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1389  
Db 469 CCAGAGATTGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 528  
  
QY 1390 GTCTGCTTCCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 1449  
Db 529 GTCTGCTTCCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 587  
  
QY 1450 CAAAAGTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGACACTTTCATCTCTCAAGC 1509  
Db 588 CAAAAGTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGACACTTTCATCTCTCAAGC 647  
  
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCC 1553  
Db 648 TCAGGGCCCATCCCTTCTCTGACG-TCTGACCCCAATTTAGTCC 690
```

```
RESULT 7  
BM842155  
LOCUS  
DEFINITION K-EST0119498 S12SNU216 Homo sapiens cDNA clone S12SNU216-56-H11 5',  
mRNA sequence.  
ACCESSION BM842155  
VERSION BM842155.1 GI:19198564  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 666)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology
```

```
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 56 row: H column: 11  
High quality sequence stop: 666.  
FEATURES  
Location/Qualifiers  
1..666  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="S12SNU216-56-H11"  
/sex="F"  
/tissue_type="Lymph node"  
/cell_type="Epithelial"  
/cell_line="SNU-216"  
/lab_host="Top10F"  
/clone_lib="S12SNU216"  
/note="Organ: Stomach; Vector: PCNS; Site_1: EcoRI;  
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then deapped  
with tobacco acid pyrophosphatase (TAP). The deapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."
```

ORIGIN

```
Query Match 27.5%; Score 431.8; DB 4; Length 666;  
Best Local Similarity 99.5%; Pred. No. 2.7e-57;  
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1090 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAGACTCGTCTCTGG 1149  
Db 232 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAGACTCGTCTCTGG 291  
  
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 1209  
Db 292 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC 351  
  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTTCATCCAGGACTCAGCAC 1269  
Db 352 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTTCATCCAGGACTCAGCAC 411  
  
QY 1270 CCGGCATCCCACTGCTGCGAGGAGACCGCTGACACTCTCTTTCAGACCCCTCATCTCTTC 1329  
Db 412 CCGGCATCCCACTGCTGCGAGGAGACCGCTGACACTCTCTTTCAGACCCCTCATCTCTTC 471  
  
QY 1330 CCAGAGATTGTGAGAAATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTCTGGAAGTTC 1389  
Db 472 CCAGAGATTGTGAGAAATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTCTGGAAGTTC 531  
  
QY 1390 GTCTGCTTCCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 1449  
Db 532 GTCTGCTTCCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 591  
  
QY 1450 CAAAAGTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGACACTTTCATCTCTCAAGC 1509  
Db 592 CAAAAGTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGACACTTTCATCTCTCAAGC 651  
  
QY 1510 TCAGGGCCCATCCCT 1524  
Db 652 TCAGGGCCCATCCCT 666
```

```

RESULT 8
AI002163/c
LOCUS      645 bp      mRNA      linear      EST 27-AUG-1998
DEFINITION o:42405.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:1619481
            3', similar to FR:Q61955 Q61955 SERINE PROTEASE INHIBITOR 5 ; , mRNA
            sequence.
ACCESSION  AI002163
VERSION     AI002163.1 GI:3202497
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 645)
AUTHORS   NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CCAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1042 Std Error: 0.00
            Seq primer: -40ml3 fwd. ET from Amersham
            High quality sequence stop: 303.

FEATURES             source
            1..645
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1619481"
            /sex="male"
            /lab_host="DH10B"
            /clone_lib="Soares testis NHT"
            /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
            polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from Clontech
            Laboratories, Inc., and primed with a Not I - oligo(dT)
            primer [5',
            TGTTACCACTCTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT7T3 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match      26.5%; Score 417; DB 1; Length 645;
Best Local Similarity 94.9%; Pred. No. 5.3e-55;
Matches 464; Conservative 0; Mismatches 20; Indels 5; Gaps 3;

QY 1084 CTCTGCAGGCTGATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTG 1143
Db 499 CTCTGCAGGCTGATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGCTCGTGT- 441

QY 1144 TCCTGGGGAGATTACCCCTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGC 1203
Db 440 --CCTGGGAGATTACCTTGT-CCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGC 384

QY 1204 AGTTTACCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCGAGTCATCCAGGACTC 1263
Db 383 AAGTTTACCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCGAGTCATCCAGGACTC 324

QY 1264 AGCACACCGGCATCCCACTGCTGTCAGGGACAGCCCTGACACTCCTTTTCAGACCCCTCAT 1323
Db 323 AGCACACCGGCATCCCACTGCTGTCAGGGACAGCCCTGACACTCCTTTTCAGACCCCTCAT 264

QY 1324 TCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCCTGGA 1383

```

```

Db 263 TCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCCTGGA 204
QY 1384 CTCAGGGTCTGCTTCCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAC 1443
Db 203 CTCAGGGTCTGCTTCCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAC 144
QY 1444 AATTTCCAAAACATGTCACAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCGACTTTCATCC 1503
Db 143 AATTTCCAAAACATGTCACAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCGACTTTCATCC 84
QY 1504 TCAAGCTCAGGGCCCATCCCTTCTCTGACGCTCGACCCCAAAATTTAGT-CCCAGAAATAA 1562
Db 83 TCAAGCTCAGGGCCCATCCCTTCTCTGACGCTCGACCCCAAAATTTAGTCCCCAGAAATAA 24
QY 1563 ACTGAGAAG 1571
Db 23 ACTGAGAAG 15

RESULT 9
LOCUS      BE898385
DEFINITION 601691219F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3951365 S',
            mRNA sequence.
ACCESSION  BE898385
VERSION     BE898385.1 GI:10364809
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 677)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTp
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCW819 row: f column: 06
            High quality sequence stop: 677.

FEATURES             source
            1..677
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3951365"
            /tissue_type="adenocarcinoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 9"
            /notes="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies). "

ORIGIN
Query Match      24.4%; Score 383.6; DB 2; Length 677;
Best Local Similarity 98.5%; Pred. No. 7.8e-50;
Matches 397; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1090 CAGGGTGATTCTGGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 1149
Db 276 CAGGGTGATTCTGGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 335

```

QY 1150 GGAGATTACCTTGTGCGCCGCGCCCAACAGACCGGGTGTCTACAGCAACTCTGCAAGTTC 1209
|||||
Db 336 GGAGATTACCTTGTGCGCCGCGCCCAACAGACCGGGTGTCTACAGCAACTCTGCAAGTTC 395
|||||
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCTGAGTCACTCCAGGACTCAGCAC 1269
|||||
Db 396 ACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCTGAGTCACTCCAGGACTCAGCAC 455
|||||
QY 1270 CCGGCATCCCACTGTCTGAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329
|||||
Db 456 CCGGCATCCCACTGTCTGAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 515
|||||
QY 1330 CCAGAGATGTTGAGAAATTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 1389
|||||
Db 516 CCAGAGATGTTGAGAAATTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 575
|||||
QY 1390 GTCTGTTCCCCCACTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAATTTTC 1449
|||||
Db 576 GNTCTGTTCCCCCACTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAATTTTC 635
|||||
QY 1450 CAAACTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGG 1492
|||||
Db 636 CAAACTGTCCA-GCGCGGGGTTCGGTCTCAATCTCCCTGGGG 677
|||||

RESULT 10

N80762/c
LOCUS
DEFINITION N80762 453 bp mRNA linear EST 29-MAR-1996
za98f06.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone
IMAGE:300611.3', mRNA sequence.

ACCESSION N80762.1 GI:1243463

VERSION N80762

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 453)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfs, T., Soares, M., Tan, F.,
Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: ml3 -40 forward

High quality sequence stop: 319.

FEATURES

source

1. .453

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:124535"

/db_xref="taxon:9606"

/clones="IMAGE:300611"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares fetal lung NbHL19W"

/note="Organ: lung; Vector: pF73D (Pharmacia) with a

modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGAGCGGCGCAATTTTTTTTTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pF73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library. Soares
fetal heart NbHL19W."

ORIGIN

Query Match 24.3%; Score 381.8; DB 7; Length 453;
Best Local Similarity 96.6%; Pred. No. 1.7e-49;
Matches 422; Conservative 0; Mismatches 9; Indels 6; Gaps 3;

QY 1141 GTGTCTCTGGGAGATTACCTTGTGCGCGGCCCAACAGA-----CCGGGTGTCTACAGAA 1196
|||||

Db 441 GTGTCTTGGNAGATTACCTTGTGCGCGGCCCAACAGAAGCGGGGTGTCTACAGAA 382
|||||

QY 1197 CTTCTGCAAGTTCACC-AGTGGATCCAGAAACCATCCAGGCCCAACTCTCTGAGTCAATCC 1255
|||||

Db 381 CTTCTGCAAGTTCACC-AGTGGATCCAGAAACCATCCAGGCCCAACTCTCTGAGTCAATCC 322
|||||

QY 1256 CAGGACTCAGCACACCGGCATCCCACTGC-TGCAGGGACAGCCCTGACACTCTCTTCA 1314
|||||

Db 321 CAGGACTCAGCACACCGGCATCCCACTGC-TGCAGGGACAGCCCTGACACTCTCTTCA 262
|||||

QY 1315 GAGCCTCATTCCTTCCAGAGATGTTGAGAAATTTCACTCTCCAGCCCTGACCCCATG 1374
|||||

Db 261 GAGCCTCATTCCTTCCAGAGATGTTGAGAAATTTCACTCTCCAGCCCTGACCCCATG 202
|||||

QY 1375 TCTCTGGACTCAGGGTCTGCTTCCCCCAACATTTGGGCTGACCGTGTCTCTAGTGAAC 1434
|||||

Db 201 TCTCTGGACTCAGGGTCTGCTTCCCCCAACATTTGGGCTGACCGTGTCTCTAGTGAAC 142
|||||

QY 1435 CTTGGGAACAATTTCCAAAACCTGTCAGGGGGGGTGGTCTCAATCTCCCTGGGGCA 1494
|||||

Db 141 CTTGGGAACAATTTCCAAAACCTGTCAGGGGGGGTGGTCTCAATCTCCCTGGGGCA 82
|||||

QY 1495 CTTTCATCTCTCAAGCTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCC 1554
|||||

Db 81 CTTTCATCTCTCAAGCTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCC 22
|||||

QY 1555 AGAAATAAATCTGAGAG 1571
|||||

Db 21 AGAAATAAATCTGAGAG 5
|||||

RESULT 11

AW105502/c

LOCUS

DEFINITION

AW105502

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

363 bp mRNA linear EST 20-OCT-1999

xd53g06.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2597530.3',

mRNA sequence.

AW105502

AW105502.1 GI:6076237

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 363)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-x@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found

Seq primer: -40UP from Gibco

FEATURES source	High quality sequence stop: 362.		FEATURES source	High quality sequence stop: 590.	
	1. 363	Location/Qualifiers		1. 590	Location/Qualifiers
	/organism="Homo sapiens"			/organism="Homo sapiens"	
	/mol_type="mRNA"			/mol_type="mRNA"	
	/db_xref="taxon:9606"			/db_xref="taxon:9606"	
	/clone="IMAGS:2597530"			/clone="S12SNU216-39-G08"	
	/tissue_type="tumor, 5 pooled (see description)"			/sex="F"	
	/lab_host="DH10B"			/tissue_type="Lymph node"	
	/clone_lib="NCI CGAP Ov23"			/cell_type="Epithelial"	
	/notes="Organ: ovary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"			/lab_host="Top10F"	
				/clone_lib="S12SNU216"	
ORIGIN				/notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."	
	Query Match 22.6%; Score 355.8; DB 2; Length 363;			Query Match 22.2%; Score 348; DB 4; Length 590;	
	Best Local Similarity 99.4%; Pred. No. 1.9e-45;			Best Local Similarity 100.0%; Pred. No. 2.7e-44;	
	Matches 357; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1213	AAGTGGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACACG 1272	QY	1090	CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGAGACTCGTGTCTGG 1149
Db	363	AAGTGGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACACG 304	Db	243	CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGAGACTCGTGTCTGG 302
QY	1273	GCATCCCCACCTGCTGAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTCCCA 1332	QY	1150	GGAGATTACCTTGTGTCGGCCGACACAGACGGGGTGTCTACAGCACTTCGCAAGTTC 1209
Db	303	GCATCCCCACCTGCTGAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTCCCA 244	Db	303	GGAGATTACCTTGTGTCGGCCGACACAGACGGGGTGTCTACAGCACTTCGCAAGTTC 362
QY	1333	GAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCAATGCTCCTGGACTCAGGGTC 1392	QY	1210	ACCAAGTGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACA 1269
Db	243	GAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCAATGCTCCTGGACTCAGGGTC 184	Db	363	ACCAAGTGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACA 422
QY	1393	TGCTTCCCCACATGGGGTGACGGTGTCTCTAGTTGAACCTCGGGAACAATTTCCAA 1452	QY	1270	CCGGCATCCCCACCTGCTCAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1329
Db	183	TGCTTCCCCACATGGGGTGACGGTGTCTCTAGTTGAACCTCGGGAACAATTTCCAA 124	Db	423	CCGGCATCCCCACCTGCTCAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 482
QY	1453	AACGTGTCAGGGGGGGTGGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGTCA 1512	QY	1330	CCAGAGATGTTGAGATGTTTCATCTCCAGCCCTGAGCCCATGCTCTCTGGACTCAGG 1389
Db	123	AACGTGTCAGGGGGGGTGGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGTCA 64	Db	483	CCAGAGATGTTGAGATGTTTCATCTCCAGCCCTGAGCCCATGCTCTCTGGACTCAGG 542
QY	1513	GGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCCAAGAAATAAAGTGAAG 1571	QY	1390	GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCT 1437
Db	63	GGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCCAAGAAATAAAGTGAAG 5	Db	543	GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCT 590
RESULT 12			RESULT 13		
BM840511			BM838406		
LOCUS			LOCUS		
DEFINITION	K-EST0117576 S12SNU216 Homo sapiens cDNA clone S12SNU216-39-G08 5', mRNA linear EST 06-MAR-2002		DEFINITION	K-EST0114765 S12SNU216 Homo sapiens cDNA clone S12SNU216-36-A03 5', mRNA linear EST 06-MAR-2002	
ACCESSION	BM840511		ACCESSION	BM838406	
VERSION	BM840511.1 GI:19196920		VERSION	BM838406.1 GI:19194815	
KEYWORDS	EST.		KEYWORDS	EST.	
SOURCE	Homo sapiens (human)		SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens		ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE			REFERENCE		
AUTHORS	Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.		AUTHORS	Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.	
TITLE	21C Frontier Korean EST Project 2001		TITLE	21C Frontier Korean EST Project 2001	
JOURNAL	Unpublished (2002)		JOURNAL	Unpublished (2002)	
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsun@mail.kribb.re.kr Plate: 39 row: G column: 08		COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsun@mail.kribb.re.kr Plate: 39 row: G column: 08	

Db 179 CACATTGGGCTGACCGGTCTCTCTAGTTGAACCTGGGAACAATTTCCAAAACGTGCCA 120
QY 1462 GGGCGGGGGTGGCTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAGCTCAGGGCCCATC 1521
Db 119 GGGCGGGGGTGGCTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAGCTCAGGGCCCATC 60
QY 1522 CCTTCTCTGACGCTCTGACCCAAATTTAGTCCCGAGAAATAAATGAGAAG 1571
Db 59 CCTTCTCTGACGCTCTGACCCAAATTTAGTCCCGAGAAATAAATGAGAAG 10

RESULT 15
W68496
LOCUS
DEFINITION
z34h08.r1 Soares fetal_heart_NbHH19W Homo sapiens cDNA clone
IMAGE:342591.5, mRNA sequence.
ACCESSION
W68496
VERSION
W68496.1 GI:1377366
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 388)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Hollman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 400 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 256.

FEATURES
source
1..388
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1267966"
/db_xref="taxon:9606"
/clone="IMAGE:342591"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DHI0B (ampicillin resistant)"
/clone_lib="Soares fetal heart_NbHH19W"
/notes="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCATCTGAAGTGGGAGCGGGCATCTTTTCTTTTCTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."

ORIGIN
Query Match 20.2%; Score 317; DB 7; Length 388;
Best Local Similarity 99.4%; Pred.No. 1.9e-39;
Matches 339; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1224 GGAAACCATCCAGGCCCAACTCTCTAGTCAATCCAGGACTCAGCACACCGGCATCCCCACC 1283

Db 1 GGAAACCATCCAGGCCCAACTCTCTAGTCAATCCAGGACTCAGCACACCGGCATCCCCACC 60
QY 1284 TGCTGAGGAGACAGCCCTGACACTCTCTTTTCAGACCCCTCATCTCTCCAGAGATGTTGAG 1343
Db 61 TGTGTGAGGAGACAGCCCTGACACTCTCTTTTCAGACCCCTCATCTCTCCAGAGATGTTGAG 120
QY 1344 AATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGGGTCTGCTTCCGCCCA 1403
Db 121 AATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGGGTCTGCTTCCGCCCA 180
QY 1404 CATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTTCCAAAACGTGCCAGG 1463
Db 181 CATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTTCCAAAACGTGCCAGG 240
QY 1464 GCGGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGCTCAGGGGCC-ATCC 1522
Db 241 GCGGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGCTCAGGGGCCAATCC 300
QY 1523 CTTTCTCTGC-AGCTCTGACCCAAATTTAGTCCCGAGAAATAA 1562
Db 301 CTTTCTCTGCAAGCTCTGACCCAAATTTAGTCCCGAGAAATAA 341

Search completed: July 4, 2005, 15:07:19
Job time : 3420.96 secs

Db 181 ACAGAGCGCTGGGACACAGGCACACAGAGTCAGAGAGAAAAGAGAGATAGAGAAAG 240

Qy	241	ACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCAGAGTACACGC	300
Db	241		
Db	241	ACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCAGAGTACACGC	300
Qy	301	AAAGGGCAGAAAGCACACAGTTTTTCAGGGTGGTGTCATATGATCATCTTCTTTTTTTTTTTTT	360
Db	301		
Qy	361	TTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTCTGCGCCAGGCTGGAGTGCAGTGGCGG	420
Db	361		
Qy	421	GATCTCGGCTCACTGCAAGCTCCGCCCTCCCGGGTTACGGCAITTTCTTCCCTCAGCCCTC	480
Db	421		
Qy	481	CCAAGTAGCTGGGACTACAGGGCGCGGCACCTAGCCCGGGTAAATTTTTTTTGTATTTT	540
Db	481		
Qy	541	GTAGAGACGGGGTTTTTCAACGTTTTTAGCGGGATCGCCTCGATCTCTCGACCTCGTGATCC	600
Db	541		
Qy	601	GCCCGCCTCGGCCCTCCAAAGTGTGGGATTTACAGGGCTGAGCCACCGCGCCGGCCCATG	660
Db	601		
Qy	661	ATCATCTTTGACTATGCTGATGTGACAGTACTTAAAGCCATCAGACTCTACCCCTTTA	720
Db	661		
Qy	721	AATATGAGTTTTGGCCAGGACCGTGGCTCATCGCTGTAATTTCCAGCACTTTGGGAGGC	780
Db	721		
Qy	781	AGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACAGCCCTGGGCCAACTGGTGA	840
Db	781		
Qy	841	CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	900
Db	841		
Qy	901	ACCTGTAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCATTGAAACCTTGGAGGCG	960
Db	901		
Qy	961	GAGGTTGAGTGGGCCGAGATCAATCACCGCCCTCCAGCTGGGCCACAGAGCAAGACT	1020
Db	961		
Qy	1021	CTGTCTCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1080
Db	1021		
Qy	1081	AAAAAATAATGCTGTCAACAAATAGAGCAGAGTGAATAAAGGAAAAATAATGGGCCA	1140
Db	1081		
Qy	1141	AGAACTCTAAGGTATTTGACAAATCATTCAGAACCTTTTAAAAAAGAAAGAAATCAAGA	1200
Db	1141		
Qy	1201	GGCATAGAAACACAGGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAGAACAA	1260
Db	1201		
Qy	1261	ACAAGGCTCTTAAGACACAGAGGAGAGAGAGAGAGTGTAGTGAGACACAGACAGAG	1320
Db	1261		

Qy	1321	AAAAAGACAGAGAGAGAGAGACAGAGAGAGAGAGACAGAGAGCGGAGGGATAGAAAGA	1380
Db	1321	AAAAAGACAGAGAGAGAGACAGAGAGACAGAGAGACAGAGAGCGGAGGGATAGAAAGA	1380
Qy	1381	GAGAGAGGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCTCAGAAGATAGCCGAGGGA	1440
Db	1381	GAGAGAGGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCTCAGAAGATAGCCGAGGGA	1440
Qy	1441	GAACCACAGAGAGATGGAAGAAGACTCTGAGAAAAAACACAGAGACAAGATGAAAAGAGG	1500
Db	1441	GAACCACAGAGAGATGGAAGAAGACTCTGAGAAAAAACACAGAGACAAGATGAAAAGAGG	1500
Qy	1501	AGTATCGAGGGTGAAACAGACAGTGGTGGAAATGACAAAATGCAGAGAGAAAAGCAAGCAA	1560
Db	1501	AGTATCGAGGGTGAAACAGACAGTGGTGGAAATGACAAAATGCAGAGAGAAAAGCAAGCAA	1560
Qy	1561	TCCAGGGCCCAAGNATGTAACCCAGAGTTGGTGAGNAGCCAGATCCTTHAAGGCTGGGG	1620
Db	1561	TCCAGGGCCCAAGNATGTAACCCAGAGTTGGTGAGNAGCCAGATCCTTHAAGGCTGGGG	1620
Qy	1621	AGGCAGGGAAGGGCTGGCCCTGGCTTCGAGAGACCCCTCCCACATTCTCCGGGCCCAGGGAG	1680
Db	1621	AGGCAGGGAAGGGCTGGCCCTGGCTTCGAGAGACCCCTCCCACATTCTCCGGGCCCAGGGAG	1680
Qy	1681	GTAGGGAGTGACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGCA	1740
Db	1681	GTAGGGAGTGACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGCA	1740
Qy	1741	GGAGGAGCTATTGCTAAGGCCCCGATAGGCACCTCATTTGCCCGGGNAATGTGCCCCAGGGAG	1800
Db	1741	GGAGGAGCTATTGCTAAGGCCCCGATAGGCACCTCATTTGCCCGGGNAATGTGCCCCAGGGAG	1800
Qy	1801	CAGTGGGTGGTTATACTCAGGCCCGGTGCCCAGAGCCCCAGAGGAGGSCAGTGGCCAGGA	1860
Db	1801	CAGTGGGTGGTTATACTCAGGCCCGGTGCCCAGAGCCCCAGAGGAGGSCAGTGGCCAGGA	1860
Qy	1861	AGGCACAGGGCCTGAGAAGTCTGCGGTGAGCTGGGAGCAAAATCCCCACCCCCCTACCTGG	1920
Db	1861	AGGCACAGGGCCTGAGAAGTCTGCGGTGAGCTGGGAGCAAAATCCCCACCCCCCTACCTGG	1920
Qy	1921	GGGACAGGGCAAGTGAGACCTGCTGAGGGTGGCTCAGCAGGCGAGGNAAGAGAGAGTGTCT	1980
Db	1921	GGGACAGGGCAAGTGAGACCTGCTGAGGGTGGCTCAGCAGGCGAGGNAAGAGAGAGTGTCT	1980
Qy	1981	GTGGTCTCTGCACCCACATCTTTCTCTGTGCCCTCTCTTGGCCCTGTCTGAGGGTGTCTAGA	2040
Db	1981	GTGGTCTCTGCACCCACATCTTTCTCTGTGCCCTCTCTTGGCCCTGTCTGAGGGTGTCTAGA	2040
Qy	2041	CTCCTATCTTTCTGAAATTTCTATAGTGGCTGGGTCTCAGCGCAGTGCAGATGGTGGCCGCTC	2100
Db	2041	CTCCTATCTTTCTGAAATTTCTATAGTGGCTGGGTCTCAGCGCAGTGCAGATGGTGGCCGCTC	2100
Qy	2101	CTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTAGGGGAGGGAGGGGAAGTGGGTTAAGG	2160
Db	2101	CTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTAGGGGAGGGAGGGGAAGTGGGTTAAGG	2160
Qy	2161	GCCTCCCGGATCGCTGGGCCTCCCAACCTCTGACATTCGCCCATCCAGGTGCAGCGGCC	2220
Db	2161	GCCTCCCGGATCGCTGGGCCTCCCAACCTCTGACATTCGCCCATCCAGGTGCAGCGGCC	2220
Qy	2221	ATGGCTACAGCAAGACCCCCCTGGATGGGTGTCTGTGCTCTGTATCACAGCCTTGCTTT	2280
Db	2221	ATGGCTACAGCAAGACCCCCCTGGATGGGTGTCTGTGCTCTGTATCACAGCCTTGCTTT	2280
Qy	2281	CTGGGGGTCA CAGGTAAACAGAACTCTGGGGTGGGAGGGTGTGGNATTCGGAGGACTGT	2340
Db	2281	CTGGGGGTCA CAGGTAAACAGAACTCTGGGGTGGGAGGGTGTGGNATTCGGAGGACTGT	2340
Qy	2341	CTCTGGGGCACTAGAGCGCTGTCCCTCTGGGAACTGTGTGAGCCTGGGCATGACTCCGG	2400
Db	2341	CTCTGGGGCACTAGAGCGCTGTCCCTCTGGGAACTGTGTGAGCCTGGGCATGACTCCGG	2400
Qy	2401	GACCGGTGAATGTGAGTCTCTCTGTGTACTTGTGGTGTGGCATGTGATGTGGCCCTGT	2460

QY 4621 CTCTTATTTCTCCAGGCCCTGCCCTGCCCCCTCAGCATGTCAGACACCCACCCCTCTAGCT 4680
Db 4621 CTCTTATTTCTCCAGGCCCTGCCCTGCCCCCTCAGCATGTCAGACACCCACCCCTCTAGCT 4680
QY 4681 GGTCTGGCCTCTTGTAGTCTGAACCCACCCCGCCAGCCCGCCCTCTGAGCCCGCC 4740
Db 4681 GGTCTGGCCTCTTGTAGTCTGAACCCACCCCGCCAGCCCGCCCTCTGAGCCCGCC 4740
QY 4741 CAACCCATTTTCGTTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTCTGTGACACC 4800
Db 4741 CAACCCATTTTCGTTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTCTGTGACACC 4800
QY 4801 CCTCTAACACCGTCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGGCCGGGGAAGACG 4860
Db 4801 CCTCTAACACCGTCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGGCCGGGGAAGACG 4860
QY 4861 CCGGTGCGATGACAGAGAGCGCGCATCATCAATGATCCGACTGCGATATGCACACC 4920
Db 4861 CCGGTGCGATGACAGAGAGCGCGCATCATCAATGATCCGACTGCGATATGCACACC 4920
QY 4921 AGCGGTGGAGCGCGCGCTCTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980
Db 4921 AGCGGTGGAGCGCGCGCTCTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980
QY 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 2
CQ874885
LOCUS CQ874885 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 3 from Patent WO2004077060.
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1
AUTHORS Diamandis, E.P. and Petraki, C.
TITLE Assay for detection of renal cell carcinoma
JOURNAL Patent: WO 2004077060-A 3 10-SEP-2004;
Mount Sinai Hospital (CA)
FEATURES
source
1. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCAGAGTGAAGCAAGAGAGGTTGAGAGCTCCCTCTGCAAGTGGCTTGATC 60
Db 1 GGGCCAGAGTGAAGCAAGAGAGGTTGAGAGCTCCCTCTGCAAGTGGCTTGATC 60
QY 61 TCCCTGCTTAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTGCTTAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAACACTACAGAAACACAGAGAGAAC 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAACACTACAGAAACACAGAGAGAAC 180
QY 181 ACAGAGAGCTGGACACAGGGACACACAGAGTACAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGACACAGGGACACACAGAGTACAGAGAGAGAGAGAGAGAGAG 240

QY 241 ACACAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACACG 300
Db 241 ACACAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACACG 300
QY 301 AAAGGGGCAAGACACAGTTTTTCAGGGTGGTGTATGATCATCTCTTTTTTTTTTTT 360
Db 301 AAAGGGGCAAGACACAGTTTTTCAGGGTGGTGTATGATCATCTCTTTTTTTTTTTT 360
QY 361 TTTTTTTTTTTTTTTTTCAGAGCGAGTCTCGCTCTGTGCCCCAGGCTGGAGTGCAGTGGCG 420
Db 361 TTTTTTTTTTTTTTTTTCAGAGCGAGTCTCGCTCTGTGCCCCAGGCTGGAGTGCAGTGGCG 420
QY 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTTCAAGCCATTCTCTCGCTCAGCCTC 480
Db 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTTCAAGCCATTCTCTCGCTCAGCCTC 480
QY 481 CCAAGTAGCTGGGACTACAGGGCCGCCCACTACAGCCCGGCTAAATTTTTTGTATTTTA 540
Db 481 CCAAGTAGCTGGGACTACAGGGCCGCCCACTACAGCCCGGCTAAATTTTTTGTATTTTA 540
QY 541 GTAGAGACGGGTTTCAACCGTTTTAGCCGGATGGCTCGATCTCCTGACCTCGTGATCC 600
Db 541 GTAGAGACGGGTTTCAACCGTTTTAGCCGGATGGCTCGATCTCCTGACCTCGTGATCC 600
QY 601 GCCCGCTCGGCTCCCAAAGTCTCGGATTAACAGCGTGAAGCCACCGCCCGGCCATG 660
Db 601 GCCCGCTCGGCTCCCAAAGTCTCGGATTAACAGCGTGAAGCCACCGCCCGGCCATG 660
QY 661 ATCATCTTCTGACTATGCTGATGACAAAGTCTGCAAGTACCTAAAGCCATCAGACTCTAC 720
Db 661 ATCATCTTCTGACTATGCTGATGACAAAGTCTGCAAGTACCTAAAGCCATCAGACTCTAC 720
QY 721 AATATCAGTTTGGGCGAGCCACCGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGC 780
Db 721 AATATCAGTTTGGGCGAGCCACCGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGC 780
QY 781 AGAGTGGGTGAATCACTTTGAGGCCAGGAGTTTGAAGCCAGCCTGGCCAAACATGGTGA 840
Db 781 AGAGTGGGTGAATCACTTTGAGGCCAGGAGTTTGAAGCCAGCCTGGCCAAACATGGTGA 840
QY 841 CTCGTCTTTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGCCGGTGTCTGGGGCAC 900
Db 841 CTCGTCTTTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGCCGGTGTCTGGGGCAC 900
QY 901 ACCTGTAATCCAGCTATCTGAGGCTGAGGCAAGAGTCACTTGAACCTCGAGGCG 960
Db 901 ACCTGTAATCCAGCTATCTGAGGCTGAGGCAAGAGTCACTTGAACCTCGAGGCG 960
QY 961 GAGTTGCAAGTGGCGGAGATCACTACCGCCCTCCAGCTGGGCGACAGAGCAAGACT 1020
Db 961 GAGTTGCAAGTGGCGGAGATCACTACCGCCCTCCAGCTGGGCGACAGAGCAAGACT 1020
QY 1021 CTGTCTCAATAAATAAATAAACAACAGCAAGTGTGTGTACCTTAGTTATATCT 1080
Db 1021 CTGTCTCAATAAATAAATAAACAACAGCAAGTGTGTGTACCTTAGTTATATCT 1080
QY 1081 AAAAAAAAAATGCTGTCAACAAATAGACAGAGTGAATAAAGGAAAAATAATGGGCCA 1140
Db 1081 AAAAAAAAAATGCTGTCAACAAATAGACAGAGTGAATAAAGGAAAAATAATGGGCCA 1140
QY 1141 AGAATCTTAAGGTATATTTGACAAATCATTTAGAACCTTTTAAAAAGAAAGATACAGA 1200
Db 1141 AGAATCTTAAGGTATATTTGACAAATCATTTAGAACCTTTTAAAAAGAAAGATACAGA 1200
QY 1201 GGCATAGAAAGACAGGGAGAACAGGGACAGAAACACCTGTGGCCCAAGGAGAACAA 1260
Db 1201 GGCATAGAAAGACAGGGAGAACAGGGAGACAGAAACACCTGTGGCCCAAGGAGAACAA 1260
QY 1261 ACAAGGCTCTTAAGACACAGGAGGAGAGAGAGAGTGTGTAGAGACAGACAGAG 1320
Db 1261 ACAAGGCTCTTAAGACACAGGAGGAGAGAGAGAGTGTGTAGAGACAGACAGAG 1320
QY 1321 AAAAAACAGAGAGAGAGACAGACAGAGAGACAGAGAGCGGAGGGATAGAAAGA 1380


```
QY 3541 TCGGCGCAATTGCACTCCAGCCTGGGCAACAAGAGCGAAACTCCGCTCTCGAAAGAAAAAAA 3600
DB |||||||
DB 3541 TCGGCGCAATTGCACTCCAGCCTGGGCAACAAGAGCGAAACTCCGCTCTCGAAAGAAAAAAA 3600
QY 3601 GAAAAAAGGCTAAGAACAGTGAATGGGACGGGAGGACTGATGATGGAGTGGGC 3660
DB |||||||
DB 3601 GAAAAAAGGCTAAGAACAGTGAATGGGACGGGAGGACTGATGATGGAGTGGGC 3660
QY 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGGAGATTGACAGGATTGAGAAGGCAT 3720
DB |||||||
DB 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGGAGATTGACAGGATTGAGAAGGCAT 3720
QY 3721 GTTTTCAATCTAGAAATTCAGAAACCTTAGGCCTGTCTTCCCTCCATCTGTGCCCCCTTAAG 3780
DB |||||||
DB 3721 GTTTTCAATCTAGAAATTCAGAAACCTTAGGCCTGTCTTCCCTCCATCTGTGCCCCCTTAAG 3780
QY 3781 CTGAGCCCTTCTTTCTGCTGCTGCTTTTCGAAACCTTAGTCTCGGCCCATGAGTCTTGACC 3840
DB |||||||
DB 3781 CTGAGCCCTTCTTTCTGCTGCTGCTTTTCGAAACCTTAGTCTCGGCCCATGAGTCTTGACC 3840
QY 3841 CCACCTCTTTTCCTCAACAGCCCTTAGGCCAGACTCTAGTGGACCCCGCCTTAAGGCCA 3900
DB |||||||
DB 3841 CCACCTCTTTTCCTCAACAGCCCTTAGGCCAGACTCTAGTGGACCCCGCCTTAAGGCCA 3900
QY 3901 CACCCCTTTGGGCGCAGGCTCCACCCCTATTCTGTGGGTACCTTCTTAGAAACCCCTTCAA 3960
DB |||||||
DB 3901 CACCCCTTTGGGCGCAGGCTCCACCCCTATTCTGTGGGTACCTTCTTAGAAACCCCTTCAA 3960
QY 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGAGAGTCTTGCTCTCTCCAGGCTGGAG 4020
DB |||||||
DB 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGAGAGTCTTGCTCTCTCCAGGCTGGAG 4020
QY 4021 TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
DB |||||||
DB 4021 TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
QY 4081 GCCTCCACCTCTGAGTAGCTGGGATTACAGGTGCGGCGCCACCAACCGCTGGCTGAATTTTT 4140
DB |||||||
DB 4081 GCCTCCACCTCTGAGTAGCTGGGATTACAGGTGCGGCGCCACCAACCGCTGGCTGAATTTTT 4140
QY 4141 GTGTCTTTAGTAGAGACAGGTTTTCACCTTTGGGCGAGGCTGTCTCAAACTCCCAACC 4200
DB |||||||
DB 4141 GTGTCTTTAGTAGAGACAGGTTTTCACCTTTGGGCGAGGCTGTCTCAAACTCCCAACC 4200
QY 4201 TCAGGTGATCGGCCACTCGGCTCCAGAGTCTCGGGGTTACAGAGGTGAGCCACCG 4260
DB |||||||
DB 4201 TCAGGTGATCGGCCACTCGGCTCCAGAGTCTCGGGGTTACAGAGGTGAGCCACCG 4260
QY 4261 CCCAGGCCCAAGTCAGAGCTCTTTATAGAGACTCTAAATGTAACCTGACCCCTGGCC 4320
DB |||||||
DB 4261 CCCAGGCCCAAGTCAGAGCTCTTTATAGAGACTCTAAATGTAACCTGACCCCTGGCC 4320
QY 4321 CTAAGTCAATTCAAACCCCTCTCGCTCCAGCTGACCTGACCCCACTACGTAGAGCC 4380
DB |||||||
DB 4321 CTAAGTCAATTCAAACCCCTCTCGCTCCAGCTGACCTGACCCCACTACGTAGAGCC 4380
QY 4381 TGACCCCACTCTTGAGACAGTTCATCCCTAAGCCCTGGTCTCCCTCCCATCCCCAG 4440
DB |||||||
DB 4381 TGACCCCACTCTTTGAGACAGTTCATCCCTAAGCCCTGGTCTCCCTCCCATCCCCAG 4440
QY 4441 GCTCCAGCCCAAGTCTGGCACTACCCCTGAGCTTGTCCAGGAATCTGTACCCCAAT 4500
DB |||||||
DB 4441 GCTCCAGCCCAAGTCTGGCACTACCCCTGAGCTTGTCCAGGAATCTGTACCCCAAT 4500
QY 4501 TTTACCTCAATGTATCTAGCCAAATTCAGGAATCTGTGAGGTCAGGTAGAGTCA 4560
DB |||||||
DB 4501 TTTACCTCAATGTATCTAGCCAAATTCAGGAATCTGTGAGGTCAGGTAGAGTCA 4560
QY 4561 GTAACCTTACGTAGCTGGCTGTCTTGTAGCTTTGAGCTGGGCTTGTGAGGTCGA 4620
DB |||||||
DB 4561 GTAACCTTACGTAGCTGGGCTGTCTTGTAGCTTTGAGCTGGGCTTGTGAGGTCGA 4620
```

```
QY 4621 CTCTTATTTCTCAGGCGCTGCCCCCTGCCCCCTCAGCATGTGACACCCACCCCTCTAGCT 4680
DB |||||||
DB 4621 CTCTTATTTCTCAGGCGCTGCCCCCTGCCCCCTCAGCATGTGACACCCACCCCTCTAGCT 4680
QY 4681 GGTCTGGCCTTTGAGTCTGAAACCCACCCCGGCAAGCCCGGCTCTGAGCCCCGCC 4740
DB |||||||
DB 4681 GGTCTGGCCTTTGAGTCTGAAACCCACCCCGGCAAGCCCGGCTCTGAGCCCCGCC 4740
QY 4741 CAACCCATTTTCGTTTCCAGAGCATGTTCTCCGCCAACATGATGTTTCTGTGACACC 4800
DB |||||||
DB 4741 CAACCCATTTTCGTTTCCAGAGCATGTTCTCCGCCAACATGATGTTTCTGTGACACC 4800
QY 4801 CCTCTAACACCGTCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGGCCGGGAAAGC 4860
DB |||||||
DB 4801 CCTCTAACACCGTCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGGCCGGGAAAGC 4860
QY 4861 CCCGTTGATGACAGAGCAGCGGCATCATCAATGATCCGACTGCGATATGCACACC 4920
DB |||||||
DB 4861 CCCGTTGATGACAGAGCAGCGGCATCATCAATGATCCGACTGCGATATGCACACC 4920
QY 4921 AGCCGTGGCAGCGCGCTGTTCTAAGGCCAACACGAGCTCTACTGCGGGCGGTGTGG 4980
DB |||||||
DB 4921 AGCCGTGGCAGCGCGCTGTTCTAAGGCCAACACGAGCTCTACTGCGGGCGGTGTGG 4980
QY 4981 TGCATCCACAGTGGCTGCTC 5000
DB |||||||
DB 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 3
CQ874961
LOCUS 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION CQ874961
VERSION CQ874961.1 GI:52748060
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Diamandis, E.P.
TITLE Multiple marker assay for detection of ovarian cancer
JOURNAL Patent: WO 2004075713-A 5 10-SEP-2004;
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
source 1. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGSCCCAGAGTGAAGCAAGAGAGAGTGTGAGAGTCCCTCTGCAAGTGGCTTGAGTC 60
DB |||||||
DB 1 GGSCCCAGAGTGAAGCAAGAGAGAGTGTGAGAGTCCCTCTGCAAGTGGCTTGAGTC 60
QY 61 TCCCTGCTCCCTAAATTCAGGGAGGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB |||||||
DB 61 TCCCTGCTCCCTAAATTCAGGGAGGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGATTAACAACCTACAGAAACACAGAGAGAAC 180
DB |||||||
DB 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGATTAACAACCTACAGAAACACAGAGAGAAC 180
QY 181 ACAGAGAGCTGGGACACAGGGGACACAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAG 240
DB |||||||
DB 181 ACAGAGAGCTGGGACACAGGGGACACAGAGTTCAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 ACACAAATGGAGACACAGAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCAGATAC 300
```

[illegible]


```
Db 4621 CTCCTATTTCTCCAGGCGCTGCCCCCTGCCCCCTCAGCATGTGACACACCCCTCTAGCT 4680
Qy 4681 GGTCTGGGCTCTTGAGTCTGAACCCACCCAGCCCAAGCCCGCTCTCAGCCCCGCC 4740
Db 4681 GGTCTGGGCTCTTGAGTCTGAACCCACCCAGCCCAAGCCCGCTCTCAGCCCCGCC 4740
Qy 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACAATGATGTTTCTGTGACACC 4800
Db 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACAATGATGTTTCTGTGACACC 4800
Qy 4801 CCTCTAACACCGTGCCTCTCGGAGCAACACAGGACCTGGAGCTGGGCGCGGGAAGACG 4860
Db 4801 CCTCTAACACCGTGCCTCTCGGAGCAACACAGGACCTGGAGCTGGGCGCGGGAAGACG 4860
Qy 4861 CCGGTCGGATGACAGAGCGCGCATCATCAATGATGATGATGATGATGATGATGATGATG 4920
Db 4861 CCGGTCGGATGACAGAGCGCGCATCATCAATGATGATGATGATGATGATGATGATGATG 4920
Qy 4921 AGCGTGGAGCGCGCTGTTGCTTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTGG 4980
Db 4921 AGCGTGGAGCGCGCTGTTGCTTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTGG 4980
Qy 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 4
AF135028
LOCUS
DEFINITION Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
ACCESSION AF135028
VERSION AF135028.1 GI:4589282
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Yousef, G.M. and Diamandis, E.P.
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation
JOURNAL J. Biol. Chem. 274 (53), 37511-37516 (1999)
MEDLINE 20076408
PUBMED 10608802
REFERENCE
AUTHORS Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Obierzu, C.V.
TITLE The new human kallikrein gene family: implications in carcinogenesis
JOURNAL Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
MEDLINE 21121728
PUBMED 10675891
REFERENCE
AUTHORS Yousef, G.M., Luo, L.Y. and Diamandis, E.P.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada
FEATURES
Source
1..11570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101..2331,2210..2293,4762..5023,5763..6019,6105..6238,11092..11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221..2293,4762..5023,5763..6019,6105..6238,11092..11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"
```

```
/protein_id="AAD26429.1"
/db_xref="GI:4589283"
/translation="MATARPPMMVLCALITALLGLVTEHVLANNVSDCHPSNTVPS
GNSQDLGAGEDARDSSRIINGSDDCMHTQPMOALLRLPNQYCGAVLHPQW
LLTAARCKRVFRVRLHSGSLSPVYSGQQMFQGVKSIPIHPGVSHPGSDMLIKLN
RR1RPTKDVQPIINVSHCPAGTKCLVSGWGTTSPOVHPFKVLOCLNLSVLSOKRCE
DAYPRO1DDTMFCAGDKAGRDSQGGSGFPVVCNGSLQGLVSGDYPFCARPNRGVYT
NLCKFYKWIQETIQANS"
ORIGIN
Query Match 100.0%; Score 5000; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGCCCAAGTGAAGGCAAGAGAGTGTGAGAGTCCCTCTGCAAGTGGCTTGAGTC 60
Db 1 GGGCCCAAGTGAAGGCAAGAGAGTGTGAGAGTCCCTCTGCAAGTGGCTTGAGTC 60
Qy 61 TCCCTCCCTAAATGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTCCCTAAATGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 ACACAAATGAGACACAGAGAGTGTAAAGAGAGAGAGATTAACAGAGTCCAGATACAGC 300
Db 241 ACACAAATGAGACACAGAGAGTGTAAAGAGAGAGAGATTAACAGAGTCCAGATACAGC 300
Qy 301 AAAGGGGACAGAGACAGTGTTCAGGGTGGTGTCTATGATCATCTCTCTTTTTTTTTT 360
Db 301 AAAGGGGACAGAGACAGTGTTCAGGGTGGTGTCTATGATCATCTCTCTTTTTTTTTT 360
Qy 361 TTTTCTTTTTTTTTTTGAGACGAGTCTCGTCTGTGTGCCAGGCTGAGTGCAGTGGCGG 420
Db 361 TTTTCTTTTTTTTTTTGAGACGAGTCTCGTCTGTGTGCCAGGCTGAGTGCAGTGGCGG 420
Qy 421 GATCTCGGCTCACTGCAAGCTCCGCGTTCAGGCGATTCCTGCTGCTCAGCTC 480
Db 421 GATCTCGGCTCACTGCAAGCTCCGCGTTCAGGCGATTCCTGCTGCTCAGCTC 480
Qy 481 CCAAGTAGCTGGGACTACAGCGCCCGCCACTACGCCCGGCTAATTTTTTTTGTATTTTA 540
Db 481 CCAAGTAGCTGGGACTACAGCGCCCGCCACTACGCCCGGCTAATTTTTTTTGTATTTTA 540
Qy 541 GTAGAGACGGGGTTTCAACCGTTTTAGCCGGATGGCTCGATCTCTGACCTCGTGATCC 600
Db 541 GTAGAGACGGGGTTTCAACCGTTTTAGCCGGATGGCTCGATCTCTGACCTCGTGATCC 600
Qy 601 CCGCGCTCGGCTCCCAAGTGTGGATTTACAGGCTGAGCCAGCCCGCCCGCCCATG 660
Db 601 CCGCGCTCGGCTCCCAAGTGTGGATTTACAGGCTGAGCCAGCCCGCCCGCCCATG 660
Qy 661 ATCATCTTCTTGACTATGCTGACAAAGTACTAAAGCCATCAGACTTACCTCTTTA 720
Db 661 ATCATCTTCTTGACTATGCTGACAAAGTACTAAAGCCATCAGACTTACCTCTTTA 720
Qy 721 AATATGAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAATTCAGACACTTTGGGAGGC 780
Db 721 AATATGAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAATTCAGACACTTTGGGAGGC 780
Qy 781 AGAGTGGTGAATCACTTTCAGGCGAGAGTTTGGAGACCAAGCTTGGGCAACATGGTGA 840
Db 781 AGAGTGGTGAATCACTTTCAGGCGAGAGTTTGGAGACCAAGCTTGGGCAACATGGTGA 840
Qy 841 CTCTGTCTTTTACTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 CTCTGTCTTTTACTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
```

Qy	901	ACCTGTAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAAACCTCGAGGCG	960
Dd	901	ACCTGTAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAAACCTCGAGGCG	960
Qy	961	GAGGTGAGTGGCGCGAGATCAATCAACCCCTCTCAGCCCTGGCGGACAGCAAGACT	1020
Dd	961	GAGGTGAGTGGCGCGAGATCAATCAACCCCTCTCAGCCCTGGCGGACAGCAAGACT	1020
Qy	1021	CTGTCTCAATAAATAAATAACAAACGAAACAGCAGTTTGTCTACCTTAGTTATATCT	1080
Dd	1021	CTGTCTCAATAAATAAATAACAAACGAAACAGCAGTTTGTCTACCTTAGTTATATCT	1080
Qy	1081	AAAAAATAAATGCTGTCAACAAATAGAGCAGAAAGTGAATAAAGAAAAATAAATGGGCCA	1140
Dd	1081	AAAAAATAAATGCTGTCAACAAATAGAGCAGAAAGTGAATAAAGAAAAATAAATGGGCCA	1140
Qy	1141	AGAACTCTAAGGTATATTTGACAAATCTCAGAACCTTTAAAAAGAAAGAAATCAAGA	1200
Dd	1141	AGAACTCTAAGGTATATTTGACAAATCAATCAGAACCTTTAAAAAGAAAGAAATCAAGA	1200
Qy	1201	GGCATAGAAACACAGGAGGAGAACAGGAGACAGAAACACCTGTGGCCCAAGGAAACAA	1260
Dd	1201	GGCATAGAAACACAGGAGGAGAACAGGAGACAGAAACACCTGTGGCCCAAGGAGAAACAA	1260
Qy	1261	ACAAAGGCTCTTAAGACACAGAGGAGGAGAGAGAGAGTGAAGAGACAGACAGAG	1320
Dd	1261	ACAAAGGCTCTTAAGACACAGAGGAGGAGAGAGAGAGTGAAGAGACAGACAGAG	1320
Qy	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAGGATAGAAGA	1380
Dd	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAGGATAGAAGA	1380
Qy	1381	GAGAGAGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGA	1440
Dd	1381	GAGAGAGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGA	1440
Qy	1441	GAACACAGAGAGATGGAAGAAGACTCTGAGAAAAAACACAGAGACAAAGATGGAAGAAGG	1500
Dd	1441	GAACACAGAGAGATGGAAGAAGACTCTGAGAAAAAACACAGAGACAAAGATGGAAGAAGG	1500
Qy	1501	AGTATCAGGGGTGAACACAGAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA	1560
Dd	1501	AGTATCAGGGGTGAACACAGAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA	1560
Qy	1561	TCCAGGCGCCAAAGATAGTGACCCAGAGTTGGTGAGAGCCAGATCCCTTAAGGCTGGGG	1620
Dd	1561	TCCAGGCGCCAAAGATAGTGACCCAGAGTTGGTGAGAGCCAGATCCCTTAAGGCTGGGG	1620
Qy	1621	AGGACGGGAAGGGCTGGCCCTGGCTTCCGAGAGACCCCTCCCAATCTCCGGGCCAGGGAG	1680
Dd	1621	AGGACGGGAAGGGCTGGCCCTGGCTTCCGAGAGACCCCTCCCAATCTCCGGGCCAGGGAG	1680
Qy	1681	GTAAGGAGTGACATTCGCGACTGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGCA	1740
Dd	1681	GTAAGGAGTGACATTCGCGACTGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGCA	1740
Qy	1741	GGAGGAGCTATTGCTAAGGCCCGATAGCACTCATTTGCCCGGGAATGTGCCCCAGGGAG	1800
Dd	1741	GGAGGAGCTATTGCTAAGGCCCGATAGCACTCATTTGCCCGGGAATGTGCCCCAGGGAG	1800
Qy	1801	CAGTGGGTGGTTATAAATCTCAGGCCCGGTGCCAGAGCCACAGGAGGAGCGAGTGGCCAGGA	1860
Dd	1801	CAGTGGGTGGTTATAAATCTCAGGCCCGGTGCCAGAGCCACAGGAGGAGCGAGTGGCCAGGA	1860
Qy	1861	AGGCACAGGCTGAGAAAGTCTGCGGCTGAGCTGGGAGCAATCCCCACCCCTACCTGG	1920
Dd	1861	AGGCACAGGCTGAGAAAGTCTGCGGCTGAGCTGGGAGCAATCCCCACCCCTACCTGG	1920
Qy	1921	GGGACAGGCAAGTGCAGACTGGTGAAGGTGGCTCAGCAGGACGAGAGAGGTGCT	1980
Dd	1921	GGGACAGGCAAGTGCAGACTGGTGAAGGTGGCTCAGCAGGCAAGGAGAGGTGCT	1980

Qy	1981	GTGCGTCTGACCAACACATCTTTCTGTCTCCCTCTGTGCCCTCTGTGAGGCTGTCTAGA	2040
Dd	1981	GTGCGTCTGACCAACACATCTTTCTGTCTCCCTCTGTGCCCTCTGTGAGGCTGTCTAGA	2040
Qy	2041	CTCTATCTTCTGAATCTATAGTGCCTGGGTCTCAGCGAGAGTCCGATGGTGGCCCGTC	2100
Dd	2041	CTCTATCTTCTGAATCTATAGTGCCTGGGTCTCAGCGAGAGTCCGATGGTGGCCCGTC	2100
Qy	2101	CTTGTGTTCTCTCTACCTGGGGAATAAGTGAAGGAGGAGGGAAGTGGGTTAAAG	2160
Dd	2101	CTTGTGTTCTCTCTACCTGGGGAATAAGTGAAGGAGGAGGGAAGTGGGTTAAAG	2160
Qy	2161	GCTCCCGGATCGCTGGGCTCCCAACCTCTTGACATTTCCCATCCAGGTGACGCGGCC	2220
Dd	2161	GCTCCCGGATCGCTGGGCTCCCAACCTCTTGACATTTCCCATCCAGGTGACGCGGCC	2220
Qy	2221	ATGGCTACAGCAAGACCCCTCGATGTGGGTGCTCTGTGCTCTGATCAGCCTTGCTT	2280
Dd	2221	ATGGCTACAGCAAGACCCCTCGATGTGGGTGCTCTGTGCTCTGATCAGCCTTGCTT	2280
Qy	2281	CTGGGGTCAACAGTAACCAAGAACTCTGGGGTGGAGGGTTGTGGGATTTGGAGACTGT	2340
Dd	2281	CTGGGGTCAACAGTAACCAAGAACTCTGGGGTGGAGGGTTGTGGGATTTGGAGACTGT	2340
Qy	2341	CTCTGGGCACTAGAGCGCCTGCTCCCTGGGGAACCTGTGTGAGCCTGGGCATGACTCCGG	2400
Dd	2341	CTCTGGGCACTAGAGCGCCTGCTCCCTGGGGAACCTGTGTGAGCCTGGGCATGACTCCGG	2400
Qy	2401	GACGGGTGAATGTGAGTCTCTGTCTGTACTTGTGGTTGTGGCATGTATGTGCCCTGT	2460
Dd	2401	GACGGGTGAATGTGAGTCTCTGTCTGTACTTGTGGTTGTGGCATGTATGTGCCCTGT	2460
Qy	2461	GACTGCCAGGTGTGTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCGG	2520
Dd	2461	GACTGCCAGGTGTGTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCGG	2520
Qy	2521	CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTTTGTGATTTGTGTGATTTAAG	2580
Dd	2521	CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTTTGTGATTTGTGTGATTTAAG	2580
Qy	2581	ATTGTGTGGCTCCACAGCTGTGGGTGAATGCATGTATAGCACTGGGGGTGTTCACTGT	2640
Dd	2581	ATTGTGTGGCTCCACAGCTGTGGGTGAATGCATGTATAGCACTGGGGGTGTTCACTGT	2640
Qy	2641	GTCTTTGGCTGTGTGTGACTTTGCATTATATGATATGATGATGATGATGATGATGATG	2700
Dd	2641	GTCTTTGGCTGTGTGTGACTTTGCATTATATGATATGATGATGATGATGATGATGATG	2700
Qy	2701	TCCCTGAGGTCCCGGATTCGCTGCAACAAAAGTGGTCAATCACCATGGAAGCTGTGACT	2760
Dd	2701	TCCCTGAGGTCCCGGATTCGCTGCAACAAAAGTGGTCAATCACCATGGAAGCTGTGACT	2760
Qy	2761	GTGTGCTTGTGAGGGATTTATGTGATTTGTGCTGATGTGATGATGATGATGATGATGATG	2820
Dd	2761	GTGTGCTTGTGAGGGATTTATGTGATTTGTGCTGATGTGATGATGATGATGATGATGATG	2820
Qy	2821	TTTGTGACCGTGTGACTTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG	2880
Dd	2821	TTTGTGACCGTGTGACTTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG	2880
Qy	2881	TCCTGTGAGGGCTGTAAATGCTACTGTATGTGTGATGATGATGATGATGATGATGATG	2940
Dd	2881	TCCTGTGAGGGCTGTAAATGCTACTGTATGTGTGATGATGATGATGATGATGATGATG	2940
Qy	2941	TTCTGTCTGTGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC	3000
Dd	2941	TTCTGTCTGTGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC	3000
Qy	3001	CAGGTGACTGACTTGCAGTGTGCTGTGTGAGAGAGTGTGTGAGTGTGAGTGTGAGTGTG	3060
Dd	3001	CAGGTGACTGACTTGCAGTGTGCTGTGTGAGAGAGTGTGTGAGTGTGAGTGTGAGTGTG	3060
Qy	3061	TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGSGATGAGGGTGTGCGATCCC	3120

Db 3061 TGTCACACACCGCATCTGTGCGGCACCTGAGACACTGTGATGAGGGTGTGCGATCCC 3120
Qy 3121 GCTAGGCTGCCCGGAGGGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACTGTGTG 3180
Db 3121 GCTAGGCTGCCCGGAGGGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACTGTGTG 3180
Qy 3181 GAGCGCAACATGGCGGTGTCTCAGAACTGCGGTGCGTGTGCTGGCTGTACTCTGTGTGCG 3240
Db 3181 GAGCGCAACATGGCGGTGTCTCAGAACTGCGGTGCGTGTGCTGGCTGTACTCTGTGTGCG 3240
Qy 3241 GCGTGGTCTTTGGGGTGAGTTTCGTAATGATGTGGTGCAGAGGCGCATCAGCAAGGGTAA 3300
Db 3241 GCGTGGTCTTTGGGGTGAGTTTCGTAATGATGTGGTGCAGAGGCGCATCAGCAAGGGTAA 3300
Qy 3301 GAAACAGGCGCGGCGGTGCTCAGCGCTGTATCCAGGCCCTTTGGAGGCCGAGGCA 3360
Db 3301 GAAACAGGCGCGGCGGTGCTCAGCGCTGTATCCAGGCCCTTTGGAGGCCGAGGCA 3360
Qy 3361 GCGCGATCACCCTGAGGTCGCGAGATCGAGGCCAGCTGACCAACATGGAGAACCCCGTCT 3420
Db 3361 GCGCGATCACCCTGAGGTCGCGAGATCGAGGCCAGCTGACCAACATGGAGAACCCCGTCT 3420
Qy 3421 CTACTAAAAATACAAAAATTTAGCTGTGTGTGGTGGCGGTGCTGTATCCCGAGCTACTC 3480
Db 3421 CTACTAAAAATACAAAAATTTAGCTGTGTGTGGTGGCGGTGCTGTATCCCGAGCTACTC 3480
Qy 3481 GGGAGACTGGGGCAGAAAAATTCGCTTGAAACCCGGAGGTGAGGTTGGGTGAGCCGAGA 3540
Db 3481 GGGAGACTGGGGCAGAAAAATTCGCTTGAAACCCGGAGGTGAGGTTGGGTGAGCCGAGA 3540
Qy 3541 TCGCGCATTTGCATCTCAGCGCTGGGCAACAGAGCGAACTCCGCTCGAAGAAAGAAAA 3600
Db 3541 TCGCGCATTTGCATCTCAGCGCTGGGCAACAGAGCGAACTCCGCTCGAAGAAAGAAAA 3600
Qy 3601 GAAAAAAGGGTGAAGAACAGTGAATGGGCAACAGAGGACTGATGATGGAGTGGGGC 3660
Db 3601 GAAAAAAGGGTGAAGAACAGTGAATGGGCAACAGAGGACTGATGATGGAGTGGGGC 3660
Qy 3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAGGCAT 3720
Db 3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAGGCAT 3720
Qy 3721 GTTTTCATCTAGAAATTCAGAAACCTAGGCGCTGCTTCCCTCCATGTGGCCCTTAAG 3780
Db 3721 GTTTTCATCTAGAAATTCAGAAACCTAGGCGCTGCTTCCCTCCATGTGGCCCTTAAG 3780
Qy 3781 CTGAGCCCTCTTCTCTGCTGCTGCTTTCGGAACCTAGCTCCGCCATGAGCTCTGACC 3840
Db 3781 CTGAGCCCTCTTCTCTGCTGCTGCTTTCGGAACCTAGCTCCGCCATGAGCTCTGACC 3840
Qy 3841 CCACCTCTTTTCTCAACACGCCCTAGGCCAGACTCTAGTGGACCCCGCTTAAGGCCA 3900
Db 3841 CCACCTCTTTTCTCAACACGCCCTAGGCCAGACTCTAGTGGACCCCGCTTAAGGCCA 3900
Qy 3901 CACCCCTTTGGCCAGGCTCACCCTCTATTCTGTGGTACTCTTAGAACCCCTTCAA 3960
Db 3901 CACCCCTTTGGCCAGGCTCACCCTCTATTCTGTGGTACTCTTAGAACCCCTTCAA 3960
Qy 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGCTGGAG 4020
Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGTCTCTCTCTCCAGGCTGGAG 4020
Qy 4021 TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
Db 4021 TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
Qy 4081 GCCTCCACCTCTGAGTAGCTGGGATTACAGGTGCGGCCACACCGCTGGCTAATTTT 4140
Db 4081 GCCTCCACCTCTGAGTAGCTGGGATTACAGGTGCGGCCACACCGCTGGCTAATTTT 4140
Qy 4141 GTGTCTTTAGTAGACAGGGTTTACCTTCTTGGCCAGGCTGTCTCAACTCCCAACC 4200

Db 4141 GTGTCTTTAGTAGACAGGGTTTACCTTGTGGCCAGGCTGTCTCAAACTCCCAACC 4200
Qy 4201 TCAGGTGATCGGCCACTCGGCTCCAGAGTGTCTGGGTACAGCGGTGAGCCACCGC 4260
Db 4201 TCAGGTGATCGGCCACTCGGCTCCAGAGTGTCTGGGTACAGCGGTGAGCCACCGC 4260
Qy 4261 CCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTTAAATGATGAAACCTGACCTGGCC 4320
Db 4261 CCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTTAAATGATGAAACCTGACCTGGCC 4320
Qy 4321 CTAACTAAGTCAATTTCAAAACCCCTCTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380
Db 4321 CTAACTAAGTCAATTTCAAAACCCCTCTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380
Qy 4381 TGACCCCACTCTTGAGACAGTTCATCCCTTAAGGCCCTGCTCTCCCTCCCACTCCCGAG 4440
Db 4381 TGACCCCACTCTTGAGACAGTTCATCCCTTAAGGCCCTGCTCTCCCTCCCACTCCCGAG 4440
Qy 4441 GCTCCAGCCCCACAGCTTTGGGCACTACCCCTGAGCTTGTCCAGGAATCTCTGTACCCAAT 4500
Db 4441 GCTCCAGCCCCACAGCTTTGGGCACTACCCCTGAGCTTGTCCAGGAATCTCTGTACCCAAT 4500
Qy 4501 TTTAACCTCACTGTAGTCTTAGCCAAATTCAGGAATCTGTGAGGTTCCAGTTAGAGTCCA 4560
Db 4501 TTTAACCTCACTGTAGTCTTAGCCAAATTCAGGAATCTGTGAGGTTCCAGTTAGAGTCCA 4560
Qy 4561 GTAAACCTACTGAGCTGGGCTGTCTGCTTGGAGCTTGGAGGTTGAGAGGTGCCA 4620
Db 4561 GTAAACCTACTGAGCTGGGCTGTCTGCTTGGAGCTTGGAGGTTGAGAGGTGCCA 4620
Qy 4621 CTCTTATTTCCAGGCGCTGCCCTGCCCTCAGCATGTCCAGACACCCACCTCTAGCT 4680
Db 4621 CTCTTATTTCCAGGCGCTGCCCTGCCCTCAGCATGTCCAGACACCCACCTCTAGCT 4680
Qy 4681 GGTCGSCCTCTTGTAGTCTGAAACCCACCCCAAGCCCGCTCTGAGGCCCGGCC 4740
Db 4681 GGTCGSCCTCTTGTAGTCTGAAACCCACCCCAAGCCCGCTCTGAGGCCCGGCC 4740
Qy 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCAACAATGATGTTTCTGTGACCAAC 4800
Db 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCAACAATGATGTTTCTGTGACCAAC 4800
Qy 4801 CCTCTAACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAAGACG 4860
Db 4801 CCTCTAACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAAGACG 4860
Qy 4861 CCGGTCGGATGACAGCAGCAGCGCATCATCATGATCCGACTGCGATATGCACACCC 4920
Db 4861 CCGGTCGGATGACAGCAGCAGCGCATCATCATGATCCGACTGCGATATGCACACCC 4920
Qy 4921 AGCGTGGCAGGCGCGCTGTGCTAAGGCCCAACCAAGCTCTACTGCGGGCGGTGTGG 4980
Db 4921 AGCGTGGCAGGCGCGCTGTGCTAAGGCCCAACCAAGCTCTACTGCGGGCGGTGTGG 4980
Qy 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 5

AC011483/c

LOCUS AC011483 107487 bp DNA linear PRI 01-JUL-2002
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.

ACCESSION AC011483

VERSION AC011483.7 GI:21637461

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission

QY 3840 CCCACCTCTTCTCCTCAACACCGCCCTAGGCCAGACTCTAGTGGACCCCGCCCTAAGGCC 3899
Db 56380 CCCACCTCTTCTCCTCAACACCGCCCTAGGCCAGACTCTAGTGGACCCCGCCCTAAGGCC 56321
QY 3900 ACACCCCTTTGGGCGCAGCTCCACCCCTATTCTGTGGGTACCTTCTAGAACCCCTTCA 3959
Db 56320 ACACCCCTTTGGGCGCAGCTCCACCCCTATTCTGTGGGTACCTTCTAGAACCCCTTCA 56261
QY 3960 AAGTCAGAGCTTTTCTTTTCTTTTGGAGACAGTCTGTCTCTCTCCAGGCTGGA 4019
Db 56260 AAGTCAGAGCTTTTCTTTTCTTTTGGAGACAGTCTGTCTCTCTCCAGGCTGGA 56201
QY 4020 GTGCAGTGGGTGATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATCTCG 4079
Db 56200 GTGCAGTGGGTGATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATCTCG 56141
QY 4080 TGCCTCAACCTCTCTAGTAGCTGGGATTACAGTGGCGGCCACACCGCTGGCTAAATTT 4139
Db 56140 TGCCTCAACCTCTCTAGTAGCTGGGATTACAGTGGCGGCCACACCGCTGGCTAAATTT 56081
QY 4140 TGTGCTTTAGTAGACAGAGGTTTCACTTGTGGCCAGGCTGTCTCAAACTCCCAAC 4199
Db 56080 TGTGCTTTAGTAGACAGAGGTTTCACTTGTGGCCAGGCTGTCTCAAACTCCCAAC 56021
QY 4200 CTCAAGTGATCCGCCCACTCGGCTCCAGAGTGTGGGTTACAGGCGTGAGCCACCG 4259
Db 56020 CTCAAGTGATCCGCCCACTCGGCTCCAGAGTGTGGGTTACAGGCGTGAGCCACCG 55961
QY 4260 CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAACTGTAACCTGACCTGGC 4319
Db 55960 CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAACTGTAACCTGACCTGGC 55901
QY 4320 CCTAACTAAGTCAATTCNAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGCC 4379
Db 55900 CCTAACTAAGTCAATTCNAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGCC 55841
QY 4380 CTGACCCCACTTCTTGAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCAATCCCA 4439
Db 55840 CTGACCCCACTTCTTGAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCAATCCCA 55781
QY 4440 GGCTCCAGCCCAAGCTTTGGCACTACCCCTGAGCTTTGTCAGGAATCTGTGACCCAA 4499
Db 55780 GGCTCCAGCCCAAGCTTTGGCACTACCCCTGAGCTTTGTCAGGAATCTGTGACCCAA 55721
QY 4500 TTTTACCTCACAATGATTTACCAATTCAGGAATCTGTGAGGTCAGATGAGTCC 4559
Db 55720 TTTTACCTCACAATGATTTACCAATTCAGGAATCTGTGAGGTCAGATGAGTCC 55661
QY 4560 AGTAACCTTACCTGAGCTGGGCTCTCTCTTGTAGCTTTGAGCCTGGGCTTTGAGAGGTGCC 4619
Db 55660 AGTAACCTTACCTGAGCTGGGCTCTCTCTTGTAGCTTTGAGCCTGGGCTTTGAGAGGTGCC 55601
QY 4620 ACTCTTATCTCAGGCCCTGCCCCCTGCCCTCAGCATGTGACACCCACCCCTCTAGC 4679
Db 55600 ACTCTTATCTCAGGCCCTGCCCCCTGCCCTCAGCATGTGACACACCCACCCCTCTAGC 55541
QY 4680 TGGTCTGGCTCTTGTAGTGAACCCACCCCGAGCCAGCCCGCTCTGAGCCCGCC 4739
Db 55540 TGGTCTGGCTCTTGTAGTGAACCCACCCCGAGCCAGCCCGCTCTGAGCCCGCC 55481
QY 4740 CCAACCATTTTCCGTTCCAGACGATGTTCTGCCCAACATGATGTTCTCTGTGACCCAC 4799
Db 55480 CCAACCATTTTCCGTTCCAGACGATGTTCTGCCCAACATGATGTTCTGTGACCCAC 55421
QY 4800 CCTCTAACCCCTGCTCTGAGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAGAGAC 4859
Db 55420 CCTCTAACCCCTGCTCTGAGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAGAGAC 55361
QY 4860 GCCCGGTGAGTGAACAGCAGGACCGCATCATCATGATGATCCGATGATGACACCC 4919
Db 55360 GCCCGGTGAGTGAACAGCAGGACCGCATCATCATGATGATCCGATGATGACACCC 55301

QY 4920 CAGCCGTGGCAGCCCGCTGTTGTTAAAGCCCAACAGCTCTACTGCGGGCGGTGTTG 4979
Db 55300 CAGCCGTGGCAGCCCGCTGTTGTTAAAGCCCAACAGCTCTACTGCGGGCGGTGTTG 55241
QY 4980 GTGCATCCACAGTGGCTGCTC 5000
Db 55240 GTGCATCCACAGTGGCTGCTC 55220
RESULT 6
AP243527/c
LOCUS Homo sapiens serine protease gene cluster, complete sequence. 230000 bp DNA linear PRI 21-NOV-2000
DEFINITION AF243527
ACCESSION AF243527
VERSION AF243527.1 GI:11244757
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 230000)
AUTHORS Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J.,
Moss,P., Paepel,B. and Wang,K.
TITLE Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL Gene 257 (1), 119-130 (2000)
MEDLINE 20510030
PUBMED 11054574
REFERENCE 2 (bases 1 to 230000)
AUTHORS Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J.,
Moss,P., Paepel,B. and Wang,K.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
Location/Qualifiers
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
complement(<6963..>11517)
/genes="KLK1"
complement(join(<6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/genes="KLK1"
/product="renal kallikrein"
complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/genes="KLK1"
/note="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG3353.1"
/db_xref="GI:11244758"
/translation="MWFLVLCALSLGGTGAAPQSIQRIIVGWGECQHSQPOWAAALVH
KSTPQCGILVHRQWLTAACISDNQYQLWGRHNPDDNTAQFVHVSFPHPGEN
MSLENHTQADEYDSDHMLRLKTEPADITDAVKVELFTEPEVGSFCLASGWS
IPENFSFDDLCQVDLKLIPNDEKKAHVQKVTDFMLCVGHLEGGKDTCCVDSGGPL
MCDGLVQVTSWGYVPCGTPKPSVAVRVLSYVKWIEDTIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
/product="ACO protease"
complement(join(13552..13704,14377..14530,14678..14917,
15416..15569,19204..19246))
/notes="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG3354.1"
/db_xref="GI:11244759"
/translation="MWLLITLSFLASTAAQDGLKLEBDECAHPQWQVALYERGR
FNGASLIIPHWLSAAHCRSMFVRLGEHNLKRDPQLRTTRVIPHRYEARS
HRNDIMLLRLVQPARLNPQVRPAVLTRCFHPGECACVSGWGLVPLSSPVSLDPTLHC

QY	61	TCCCTCGCCTAAATTCGAGGGAGAGGGAGCGAAGAAGACAGGGAA-GAGGAAGGGGTGGG	119
Db	142488	TCCCTCGCCTAAATTCGAGGGAGAGGGAGCGAAGAAGACAGGGAAAGGAAGGGGTGGG	142499
QY	120	GAAAGAAAGAGAGAGAGAGAGAGACAATAAACCAACTACAGAAAAACACAGAGAGAAC	179
Db	142428	GAAGAAAGAGAGAGAGAGAGAGACAATAAACCAACTACAGAAAAACACAGAGAGAAC	142369
QY	180	CACAGAGAGCCTGGGACACACAGGCACACAGAGTACAGAGAAAAGAGAAATAGAGAA	239
Db	142368	CACAGAGAGCCTGGGACACACAGGCACACAGAGTACAGAGAAAAGAGAAATAGAGAA	142309
QY	240	GACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCCAGATACAG	299
Db	142308	GACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCCAGATACAG	142249
QY	300	CAAAGGGCAGAGACACAGTTTTTAGGGTGTTCTATGATCATCTTCTTTTTTTTTTTTT	359
Db	142248	CAAAGGGCAGAGACACAGTTTTTAGGGTGTTCTATGATCATCTTCTTTTTTTTTTTTT	142189
QY	360	TTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTGCCCCAGGCTGGAGTCSAGTGGCG	419
Db	142188	TTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTGCCCCAGGCTGGAGTCSAGTGGCG	142129
QY	420	GGATCTCGGCTCACTGCAAGCTCCGCCCTCCGGGTTACGCCATTTCTCTGCTCAGCCT	479
Db	142128	GGATCTCGGCTCACTGCAAGCTCCGCCCTCCGGGTTACGCCATTTCTCTGCTCAGCCT	142069
QY	480	CCCAAGTAGCTGGAGCTACAGGCGCGGCCACTACGCCCGGGTAATTTTTTTTTTTT	539
Db	142068	CCCAAGTAGCTGGAGCTACAGGCGCGGCCACTACGCCCGGGTAATTTTTTTTTTTT	142009
QY	540	AGTAGAGACGGGTTTTACCCTTTAGCGGGATGSCCTCGATCTCTGACCTCGTGATC	599
Db	142008	AGTAGAGACGGGTTTTACCCTTTAGCGGGATGSCCTCGATCTCTGACCTCGTGATC	141949
QY	600	CGCCGCCCTCGGCCCTCCCAAAGTGTGGGATTACAGGGGTGAGCCAACGCGCGCCGCCAT	659
Db	141948	CGCCGCCCTCGGCCCTCCCAAAGTGTGGGATTACAGGGGTGAGCCAACGCGCGCCGCCAT	141889
QY	660	GATCATCTTTTGACTATGCTGATGTGACAAGTACTTAAAGCCATCAGACTCTACCCCTT	719
Db	141888	GATCATCTTTTGACTATGCTGATGTGACAAGTACTTAAAGCCATCAGACTCTACCCCTT	141829
QY	720	AAATATGAGTTTGGGCGAGCACCGTGGCTCATCGCTGTAAATCCAGCACCTTTGGGAGG	779
Db	141828	AAATATGAGTTTGGGCGAGCACCGTGGCTCATCGCTGTAAATCCAGCACCTTTGGGAGG	141769
QY	780	CAGAGGTGGGTGAATCACTTGAGGCGAGAGTTTGAGACCAGCCTGGCCCAACATGGTGAA	839
Db	141768	CAGAGGTGGGTGAATCACTTGAGGCGAGAGTTTGAGACCAGCCTGGCCCAACATGGTGAA	141709
QY	840	ACTCTGTCTTACTTAAAAAATAAAAAAAAAAAAAAAAAATCAGCGGGTGTCTGGGGCA	899
Db	141708	ACTCTGTCTTACTTAAAAAATAAAAAAAAAAAAAAAAAATCAGCGGGTGTCTGGGGCA	141649
QY	900	CACCTGTAAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAAACCTCGAGGC	959
Db	141648	CACCTGTAAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAAACCTCGAGGC	141589
QY	960	GGAGGTTGAGTGGGCCGAGATCACATCAGCCCTCCAGCCTGGGCGACAGCAAGAC	1019
Db	141588	GGAGGTTGAGTGGGCCGAGATCACATCAGCCCTCCAGCCTGGGCGACAGCAAGAC	141529
QY	1020	TCTGTCTCAAAATAAATAAACAAACGAAACAGCAGTTTGTGTACCTTAGTTATATC	1079
Db	141528	TCTGTCTCAAAATAAATAAACAAACGAAACAGCAGTTTGTGTACCTTAGTTATATC	141469
QY	1080	TAAAAAATAAATGCTGTCAACAAATAGAGCAGAGTGAATAAAGGAAATAAATGGGCC	1139
Db	141468	TAAAAAATAAATGCTGTCAACAAATAGAGCAGAGTGAATAAAGGAAATAAATGGGCC	141409

QY	1140	AAGAACTCTAAGGTATATTTGACAAATCATTTTCAGAACCTTTAAAAAGAAAGATCAACAG	11199
DB	141408	AAGAACTCTAAGGTATATTTGACAAATCATTTTCAGAACCTTTAAAAAGAAAGATCAACAG	
	141349	AAGAACTCTAAGGTATATTTGACAAATCATTTTCAGAACCTTTAAAAAGAAAGATCAACAG	141349
QY	1200	AGGCATAGAAAGACAGGGAGAAACAGGGAGACAGAAACACCTGTGCGCCCAAGGAGAACAA	1259
DB	141348	AGGCATAGAAAGACAGGGAGAAACAGGGAGACAGAAACACCTGTGCGCCCAAGGAGAACAA	141348
	141349	AGGCATAGAAAGACAGGGAGAAACAGGGAGACAGAAACACCTGTGCGCCCAAGGAGAACAA	141349
QY	13260	AACAAGGCTCTTAAGACAGACAGGAGGAGAGAGAGAGAGTGTGAGTGTGAGAGACAGACAGA	13319
DB	141288	AACAAGGCTCTTAAGACAGACAGGAGGAGAGAGAGAGTGTGAGTGTGAGAGACAGACAGA	141229
	141229	AACAAGGCTCTTAAGACAGACAGGAGGAGAGAGAGAGTGTGAGTGTGAGAGACAGACAGA	141229
QY	1320	GA AAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGCAGAGGGGATAGAAAG	1379
DB	141228	GA AAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGCAGAGGGGATAGAAAG	141169
	141169	GA AAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGCAGAGGGGATAGAAAG	141169
QY	1380	AGAGAGGGGTGCAGAGAGACACGAGATATTGTAGAGAGACTTCAGAAAGATAGCCGAGGG	1439
DB	141168	AGAGAGGGGTGCAGAGAGACACGAGATATTGTAGAGAGACTTCAGAAAGATAGCCGAGGG	141109
	141109	AGAGAGGGGTGCAGAGAGACACGAGATATTGTAGAGAGACTTCAGAAAGATAGCCGAGGG	141109
QY	1440	AGAACACAGAGAGATGGAGAAAGACTCTCGAGAAAAAACCCAGAGACAAAGATGGAAAAAG	1499
DB	141108	AGAACACAGAGAGATGGAGAGAGACTCTCGAGAAAAAACCCAGAGACAAAGATGGAAAAAG	141049
	141049	AGAACACAGAGAGATGGAGAGAGACTCTCGAGAAAAAACCCAGAGACAAAGATGGAAAAAG	141049
QY	1500	GAGTATCGAGGGTGAACACAGACAGTGGTGGAAATGACAAAAATGCAGAGAGAAAGCAAGCA	1559
DB	141048	GAGTATCGAGGGTGAACACAGACAGTGGTGGAAATGACAAAAATGCAGAGAGAAAGCAAGCA	140989
	140989	GAGTATCGAGGGTGAACACAGACAGTGGTGGAAATGACAAAAATGCAGAGAGAAAGCAAGCA	140989
QY	1560	ATCAGGGGCCAAGAAATAGTGACCCAGAGTTGGTGAGAAAGCCAGATCCTTAAGGCTGGGG	1619
DB	140988	ATCAGGGGCCAAGAAATAGTGACCCAGAGTTGGTGAGAAAGCCAGATCCTTAAGGCTGGGG	140929
	140929	ATCAGGGGCCAAGAAATAGTGACCCAGAGTTGGTGAGAAAGCCAGATCCTTAAGGCTGGGG	140929
QY	1620	GAGCAGGGAAGGGGCTGGCCCTGCTCCGAGACCCCTCCCATTTCTCCGGGCCAGGGA	1679
DB	140928	GAGCAGGGAAGGGGCTGGCCCTGCTCCGAGACCCCTCCCATTTCTCCGGGCCAGGGA	140869
	140869	GAGCAGGGAAGGGGCTGGCCCTGCTCCGAGACCCCTCCCATTTCTCCGGGCCAGGGA	140869
QY	1680	GGTAGGAGTGCACATTCGCGACTCGGGTGGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGC	1739
DB	140868	GGTAGGAGTGCACATTCGCGACTCGGGTGGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGC	140809
	140809	GGTAGGAGTGCACATTCGCGACTCGGGTGGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGC	140809
QY	1740	AGGAGGAGCTTATTCGTAAGGCCCGATAGGCACTCATTTGCCCGGGAATGTGCCCCAGGGA	1799
DB	140808	AGGAGGAGCTTATTCGTAAGGCCCGATAGGCACTCATTTGCCCGGGAATGTGCCCCAGGGA	140749
	140749	AGGAGGAGCTTATTCGTAAGGCCCGATAGGCACTCATTTGCCCGGGAATGTGCCCCAGGGA	140749
QY	1800	GCAGTGGTGGTTATACTCAGGCCCGGTGCCAGAGCCACAGAGAGGAGGCGAGTGCACAGG	1859
DB	140748	GCAGTGGTGGTTATACTCAGGCCCGGTGCCAGAGCCACAGAGAGGAGGCGAGTGCACAGG	140689
	140689	GCAGTGGTGGTTATACTCAGGCCCGGTGCCAGAGCCACAGAGAGGAGGCGAGTGCACAGG	140689
QY	1860	AAGGCACAGGCTCAGAGAGTCTGGGCTGTGAGTGGGAGCAATCCCCACCCCTACCTG	1919
DB	140688	AAGGCACAGGCTCAGAGAGTCTGGGCTGTGAGTGGGAGCAATCCCCACCCCTACCTG	140629
	140629	AAGGCACAGGCTCAGAGAGTCTGGGCTGTGAGTGGGAGCAATCCCCACCCCTACCTG	140629
QY	1920	GGGACACAGGCAAGTGAGACCTGTGAGGGTGGCTCAGACGCGAGGGAAGAGAGGTGTCT	1979
DB	140628	GGGACACAGGCAAGTGAGACCTGTGAGGGTGGCTCAGACGCGAGGGAAGAGAGGTGTCT	140569
	140569	GGGACACAGGCAAGTGAGACCTGTGAGGGTGGCTCAGACGCGAGGGAAGAGAGGTGTCT	140569
QY	1980	TGTGGGTCTGCACCCACATCTTCTGTGCCCTCCCTCTGCGCTCTCTGAGGCTGCTAG	2039
DB	140568	TGTGGGTCTGCACCCACATCTTCTGTGCCCTCCCTCTGCGCTCTCTGAGGCTGCTAG	140509
	140509	TGTGGGTCTGCACCCACATCTTCTGTGCCCTCCCTCTGCGCTCTCTGAGGCTGCTAG	140509
QY	2040	ACTCCTATCTTCTGAATTCATAGTGCCTGGGTCTCAGCGAGTCCGATGGTGGCCGT	2099
DB	140508	ACTCCTATCTTCTGAATTCATAGTGCCTGGGTCTCAGCGCAGTCCGATGGTGGCCGT	140449
	140449	ACTCCTATCTTCTGAATTCATAGTGCCTGGGTCTCAGCGCAGTCCGATGGTGGCCGT	140449
QY	2100	CCTTGTGGTTCCTCTCTACCTGGGGAATAAGGTAGGGAGGGGAGGGAAGTGGGTTAAG	2159
DB	140448	CCTTGTGGTTCCTCTCTACCTGGGGAATAAGGTAGGGAGGGGAGGGAAGTGGGTTAAG	140389
	140389	CCTTGTGGTTCCTCTCTACCTGGGGAATAAGGTAGGGAGGGGAGGGAAGTGGGTTAAG	140389
QY	2160	GGCTCCCCGGATCGCTGGGCTCCCAACCTCTGACATTCGCCATCCAGTGCAGCGCG	2219
DB	140388	GGCTCCCCGGATCGCTGGGCTCCCAACCTCTGACATTCGCCATCCAGTGCAGCGCG	140329
	140329	GGCTCCCCGGATCGCTGGGCTCCCAACCTCTGACATTCGCCATCCAGTGCAGCGCG	140329
QY	2220	CATGGCTACAGCAAGACCCCCCTGGATGTGGGTGCTCTGTGCTCTGTGATCACAGCCTTGCT	2279

Db	140328	CATGGCTACAGCAAGACCCCTTGGATGTGGGTGCTCTGTGCTCTGATCAGAGCCTTGCT	140269
Qy	2280	TCTTGGGGTCAACAGGTAAACAGAACTCTCGGGTGGAGGGTTGTGGGATTTGGAGGACTG	2339
Db	140268	TCTGGGGTCAACAGGTAAACAGAACTCTCGGGTGGAGGGTTGTGGGATTTGGAGGACTG	140209
Qy	2340	TCTCTGGGCACTAGAGCGCTGTCCCTTGGGGAACCTGTGTAGGCTTGGGCACTGCTCCG	2399
Db	140208	TCTCTGGGCACTAGAGCGCTGTCCCTTGGGGAACCTGTGTAGGCTTGGGCACTGCTCCG	140149
Qy	2400	GGACCGGTGAATGTGAGTCTCTGTCTGACTTGTGTGTGGTGTGGCATCTGTATGTGGCCCTG	2459
Db	140148	GGACCGGTGAATGTGAGTCTCTGTCTGACTTGTGTGTGGTGTGGCATCTGTATGTGGCCCTG	140089
Qy	2460	TGACTGCCAGGTGTGTGTGGGAGGGGATGCCCTTTCCCATATCAGGTGACTGTGCG	2519
Db	140088	TGACTGCCAGGTGTGTGTGGGAGGGGATGCCCTTTCCCATATCAGGTGACTGTGCG	140029
Qy	2520	CGAGGTGGCACTAACCCCTTTGAGGCTGTGTGTGGTGTGGTGTGTGTGATGTGTGCAATTA	2579
Db	140028	CGAGGTGGCACTAACCCCTTTGAGGCTGTGTGTGGTGTGGTGTGTGTGATGTGTGCAATTA	139969
Qy	2580	GATTTGTGTGGCTCCACAGCTGTGTGGGTGAATGCATGTAGCACTGGGGGTGTCACTG	2639
Db	139968	GATTTGTGTGGCTCCACAGCTGTGTGGGTGAATGCATGTAGCACTGGGGGTGTCACTG	139909
Qy	2640	TGTGTTTGGCTGTGTGTGACTTGGCATTTCTATATGACTTCAGGTATCTGCAGTTTCT	2699
Db	139908	TGTGTTTGGCTGTGTGTGACTTGGCATTTCTATATGACTTCAGGTATCTGCAGTTTCT	139849
Qy	2700	GTCCCTGAGGTCGGGATTTGGGTGCAACAAAGTGGTCAACCATGGAAGCTGTGAC	2759
Db	139848	GTCCCTGAGGTCGGGATTTGGGTGCAACAAAGTGGTCAACCATGGAAGCTGTGAC	139789
Qy	2760	TGTGTCTGCTTGCAGGCCAATATGTGATTTGGCTGAGTGTGACGTTATGGATGCCCGT	2819
Db	139788	TGTGTCTGCTTGCAGGCCAATATGTGATTTGGCTGAGTGTGACGTTATGGATGCCCGT	139729
Qy	2820	ATTTGTGACCGTGTGACTTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGT	2879
Db	139728	ATTTGTGACCGTGTGACTTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGT	139669
Qy	2880	GTCTGTGTGAGGCGGTGTAATGCTACTGTATGTGTGATGTGTGAGCTGTGTGCTGAG	2939
Db	139668	GTCTGTGTGAGGCGGTGTAATGCTACTGTATGTGTGATGTGTGAGCTGTGTGCTGAG	139609
Qy	2940	TTTCTGTCTGTGCTGGAGGATAGAGGTGACGGGTAGCTATCTCTGGAGATGGGTG	2999
Db	139608	TTTCTGTCTGTGCTGGAGGATAGAGGTGACGGGTAGCTATCTCTGGAGATGGGTG	139549
Qy	3000	CCAGGTGACTGTCTCAGTGTGTGCTGTGTGCAGAGAGATGTGCGAGCTGTGAACAT	3059
Db	139548	CCAGGTGACTGTCTCAGTGTGTGCTGTGTGCAGAGAGATGTGCGAGCTGTGAACAT	139489
Qy	3060	CTGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGTTGCGATCC	3119
Db	139488	CTGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGTTGCGATCC	139429
Qy	3120	CGCTAGGCTGCCCGGAGCGTGTGTACCTGTGGAGACAGAGCTGTATGTTAGCTGTGACCTGT	3179
Db	139428	CGCTAGGCTGCCCGGAGCGTGTGTACCTGTGGAGACAGAGCTGTATGTTAGCTGTGACCTGT	139369
Qy	3180	GGAGGCAACATGGCGTGTCTGCAGAACTCGTGGTGTGGTGTGGCTTACTGTCTGTGTG	3239
Db	139368	GGAGGCAACATGGCGTGTCTGCAGAACTCGTGGTGTGGTGTGGCTTACTGTCTGTGTG	139309
Qy	3240	CGCGTGTCTTCTGGGTGTGATTCGTAATGTGTGTGTCAGAGGCCATCAGCAAGGTA	3299
Db	139308	CGCGTGTCTTCTGGGTGTGATTCGTAATGTGTGTGTCAGAGGCCATCAGCAAGGTA	139249
Qy	3300	AGAACACGCGCGGCGGTGCTACGCTGTGTAATCCAGCCCTTTGGGAGGCGGAGGC	3359

QY	1970	GAGAGGAGTGTGTCGCTGTCGACCCACATCTTTCTCTGTCTCCCTCCTTCCCTGCTGG	2029	3021	TGTGCTGTGTGCAGAGAAGTATGTGGCAGTCTGAACATCTGTGCACACACGCGCATCTGT	3080
Db	202470	GAGAGGAGTGTGTCGCTGTCGACCCACATCTTTCTCTGTCTCCCTCCTTCCCTGCTGG	202411	201390	TGTGCTGTGTGCAGAGAAGTATGTGGCAGTCTGAACATCTGTGCACACACGCGCATCTGT	201331
QY	2030	AGGCTGTAGACTCTTCTCTGAATCTTATAGTGCCTGGGTCTCAGCGCAGTGGCGAT	2089	3081	GGGTGCACCTGAGACACTGTGGATGAGGGTGTGCGATCCGGCTAGGCTGCCCGGGAGCGT	3140
Db	202410	AGGCTGTAGACTCTTCTCTGAATCTTATAGTGCCTGGGTCTCAGCGCAGTGGCGAT	202351	201330	GGGTGCACCTGAGACACTGTGGATGAGGGTGTGCGATCCGGCTAGGCTGCCCGGGAGCGT	201271
QY	2090	GGTGGCCCGTCTGTGTGGTTCCCTCTACCTGGGGAATAAGGTAGGGGAGGGAGGGGAA	2149	3141	GTGTACTCTGGAGACAGAGCTGTATGTAGTCACCTGTGGAGGCAACATGGCGTGTCT	3200
Db	202350	GGTGGCCCGTCTGTGTGGTTCCCTCTACCTGGGGAATAAGGTAGGGGAGGGAGGGGAA	202291	201270	GTGTACTCTGGAGACAGAGCTGTATGTAGTCACCTGTGGAGGCAACATGGCGTGTCT	201211
QY	2150	GTGGGTAAAGGGTCCCGGATCGCTGGCCCTCCCAACCTCTGACATTTCCCATCCAG	2209	3201	GCAGAACTCCGTCGCTTGGCTGTACTGTCTTGTGGCGTGTCTTGTGGGTGAGT	3260
Db	202290	GTGGGTAAAGGGTCCCGGATCGCTGGCCCTCCCAACCTCTGACATTTCCCATCCAG	202231	201210	GCAGAACTCCGTCGCTTGGCTGTACTGTCTTGTGGCGTGTCTTGTGGGTGAGT	201151
QY	2210	GTGCGGGGCATGGCTACAGCAAGACCCCTTGGATGTGGGTGCTCTGTGCTGATCA	2269	3261	TGCTGAATGATGGTGTGTCAGGGCCATCAGACNAGGGTAAGAACAGGCCGGCGGTG	3320
Db	202230	GTGCGGGGCATGGCTACAGCAAGACCCCTTGGATGTGGGTGCTCTGTGCTGATCA	202171	201150	TGCTGAATGATGGTGTGTCAGGGCCATCAGCAAGGGTAAGAACAGGCCGGCGGTG	201091
QY	2270	CAGCCTTGCTTCTGGGGTCAACAGGTAAACAGAACTCTGGGTGGAGGGTGTGGGATT	2329	3321	GCTCAGCCTGTAAATCCAGCCCTTTGGGAGCCGAGGCGGATCACCTCAGGTCGG	3380
Db	202170	CAGCCTTGCTTCTGGGGTCAACAGGTAAACAGAACTCTGGGTGGAGGGTGTGGGATT	202111	201090	GCTCAGCCTGTAAATCCAGCCCTTTGGGAGCCGAGGCGGATCACCTCAGGTCGG	201031
QY	2330	GGGAGGACTGTCTCTCGGGCACTAGAGCGCTGTCCCTGGGGAACCTGTGAGCCTGGG	2389	3381	GAGATCGAGGCCAGCCGTGACCAACATGAGAGACCCCGTCTCTACTAAAAATACAAAAAT	3440
Db	202110	GGGAGGACTGTCTCTCGGGCACTAGAGCGCTGTCCCTGGGGAACCTGTGAGCCTGGG	202051	201030	GAGATCGAGGCCAGCCGTGACCAACATGAGAGACCCCGTCTCTACTAAAAATACAAAAAT	200971
QY	2390	CATGACTCGGGACCGGTGAATGTGAGTCTCTGTCTACTTGTGGTGTGCGATCGTA	2449	3441	TAGCTGGTGTGGTGGCGTGTAAATCCGAGTACTCTGGGAGACTGGGGCAGAAAAA	3500
Db	202050	CATGACTCGGGACCGGTGAATGTGAGTCTCTGTCTACTTGTGGTGTGCGATCGTA	201991	200970	TAGCTGGTGTGGTGGCGTGTGTAAATCCAGTACTCTGGGAGACTGGGGCAGAAAAA	200911
QY	2450	TGTGGCCCTGTGACTGCCAGGTGTGTGGGAGGGGATGCCCTTTCCCATATCAGG	2509	3501	TGCTTTGAACCCGGGAGGTGGAGTTGCGGTGAGCCGAGATCGCGCATTTGCACTCCAG	3560
Db	201990	TGTGGCCCTGTGACTGCCAGGTGTGTGGGAGGGGATGCCCTTTCCCATATCAGG	201931	200910	TGCTTTGAACCCGGGAGGTGGAGTTGCGGTGAGCCGAGATCGCGCATTTGCACTCCAG	200851
QY	2510	TGACTGTGCGGCAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGATTTGTGT	2569	3561	CTGGGCAACAAGAGCGAAACTCCGCTCTCGAAAGAAAAAAGAAAAAGGTAAGAA	3620
Db	201930	TGACTGTGCGGCAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGATTTGTGT	201871	200850	CTGGGCAACAAGAGCGAAACTCCGCTCTCGAAAGAAAAAAGAAAAAGGTAAGAA	200791
QY	2570	GTGCATTTAAGATTGTGTGGCTCCACAGCTGTGTGGGTGAATGATGAGCACTGGGG	2629	3621	CAAGTGAATGGGCACGGAGGACTGATGATGAGTGGGGCATGATGTAGTCTGTAGGTC	3680
Db	201870	GTGCATTTAAGATTGTGTGGCTCCACAGCTGTGTGGGTGAATGATGAGCACTGGGG	201811	200790	CAAGTGAATGGGCACGGAGGACTGATGATGAGTGGGGCATGATGTAGTCTGTAGGTC	200731
QY	2630	GTGTTCACTGTGTGTGGCTGTGTGGTGTGTGGTGTGTGATGATGATGATGATGATG	2684	3681	TGTGTGTAGAGAGAGAGATTGACAGGATTGAGAGGCAATGTTTTTCTGTGAGAAATTCAG	3740
Db	201810	GTGTTCACTGTGTGTGGCTGTGTGGTGTGTGGTGTGTGATGATGATGATGATGATG	201751	200730	TGTGTGTAGAGAGAGAGATTGACAGGATTGAGAGGCAATGTTTTTCTGTGAGAAATTCAG	200671
QY	2685	-----GTATCTGCAG-----TTCTCTCCCTGAGGTCCTGGGATG	2720	3741	AAACCTTAGGCCCTGCTCTTCCCTCCATGTCGAGGTCGAGCCCTTTCTTCTCTGCT	3800
Db	201750	ATACATATCCCTGATGATATGACAGGATACATATACATATCCCTGAGGTCCTGGGATG	201691	200670	AAACCTTAGGCCCTGCTCTTCCCTCCATGTCGAGGTCGAGCCCTTTCTTCTCTGCT	200611
QY	2721	CGTGCAACAAAGTGTGATFACACATGGAAGCTGTGACTGTGTGTGTGTGTGTGTGTGT	2780	3801	CCTGCTTTTCCGAAACCTTAGCTCGGCCCATGAGCTGTGACCCCACTCTTTTCTCAACCA	3860
Db	201690	CGTGCAACAAAGTGTGATFACACATGGAAGCTGTGACTGTGTGTGTGTGTGTGTGTGT	201631	200610	CCTGCTTTTCCGAAACCTTAGCTCGGCCCATGAGCTGTGACCCCACTCTTTTCTCAACCA	200551
QY	2781	TATGTGATTTGGCTGAGGTGACGTTATGGAATGCGCGTATTTGTGACCGTGTGACTACC	2840	3861	CGCCCTTAGCCAGACTCTAGTGGAGCCCGCTTAAAGCCACACCCCTTTGGGGCAGGCTC	3920
Db	201630	TATGTGATTTGGCTGAGGTGACGTTATGGAATGCGCGTATTTGTGACCGTGTGACTACC	201571	200550	CGCCCTTAGCCAGACTCTAGTGGAGCCCGCTTAAAGCCACACCCCTTTGGGGCAGGCTC	200491
QY	2841	TGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTGTGTGTGTGTGTGTGTGTAA	2900	3921	CACCCCTTATTCTGTGGGTACCTTCTAGAACCCCTTCAAAGTCAGAGC-TTTTTTTTTTT	3979
Db	201570	TGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTGTGTGTGTGTGTGTGTGTAA	201511	200490	CACCCCTTATTCTGTGGGTACCTTCTAGAACCCCTTCAAAGTCAGAGCTTTTTTTTTTT	200431
QY	2901	TGCTACTGTATGTGTGAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2960	3980	TTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGGTGTGAGTGTGAGTGTGATCTCGG	4039
Db	201510	TGCTACTGTATGTGTGAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	201451	200430	TTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGGTGTGAGTGTGAGTGTGATCTCGG	200371
QY	2961	ATAGAGGTGACAGGGTGTGCTCTCTGGAGATGGTGCCAGGTGACTGTGACTGTGAGTG	3020	4040	CTCAGTGCACCTCTGCGCTCCAGGTTCAAGTGAATTTCTGTGCTCCACTCTCTGAGTAG	4099
Db	201450	ATAGAGGTGACAGGGTGTGCTCTCTGGAGATGGTGCCAGGTGACTGTGACTGTGAGTG	201391	200370	CTCAGTGCACCTCTGCGCTCCAGGTTCAAGTGAATTTCTGTGCTCCACTCTCTGAGTAG	200311
				4100	CTGGGATTACAGGTGCGGCCACCGCTGGCTAATTTTTTGTGTCTTTAGTAGAGACAG	4159

Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 200792)
Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Cariga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghghi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Latic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Marquilies, E.H., Masiello, C., Maskeri, B., McDowell, J.,
Paquirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Redix-Dugue, N., Schandler, K., Schusler, M.G., Sison, C.,
Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 200792)
Green, E.D.
Direct Submission
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Groveom Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 200792)
Green, E.D.
Direct Submission
Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
Groveom Circle, Gaithersburg, MD 20877, USA
On Nov 22, 2002 this sequence version replaced gi:22218452.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dhz
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-ff
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-ff
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 2306: contig of 28306 bp in length
* 28307 28406: gap of unknown length
* 28407 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length

* 37957 73522: contig of 35566 bp in length
* 73523 73622: gap of unknown length
* 73623 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 88817: contig of 5150 bp in length
* 88818 88917: gap of unknown length
* 88918 125611: contig of 36694 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174799: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193344: contig of 6862 bp in length
* 193345 193445: gap of unknown length
* 193445 199364: contig of 5919 bp in length
* 199364 199464: gap of unknown length
* 199464 200792: contig of 1329 bp in length.

FEATURES

Source

Location/Qualifiers
1. .200792

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="CH251-355A20"

/clone_lib="CH251"

1. .28306

/note="assembly_fragment"

clone_end:Sp6

vector_side:left

28407. .37856

/notes="assembly_fragment"

37957. -73522

/note="assembly_fragment"

73623. .83567

/notes="assembly_fragment"

83668. .88817

/notes="assembly_fragment"

88918. -125611

/notes="assembly_fragment"

125712. .159879

/notes="assembly_fragment"

159980. .174698

/notes="assembly_fragment"

174799. 186382

/notes="assembly_fragment"

186483. .193344

/notes="assembly_fragment"

193445. .199363

/notes="assembly_fragment"

199464. -200792

/note="assembly_fragment"

clone_end:T7

vector_side:right

ORIGIN

Query Match 82.1%; Score 4106.4; DB 2; Length 200792;
Best Local Similarity 97.3%; Pred. No. 0; Mismatches 66; Indels 52; Gaps 7;
Matches 4268; Conservative 0;

QY 654 GGCCATGATCATCTTCTTGACTATGCTGATGACCAAGTACCTAAAGCCATCAGACTCTA 713
DB GTCTATGATCATCTTCTTGACTATGCTGATGACCAAGTACCTAAAGCCATCAGACTCTA 78438
QY 714 CCCTTTAAATATGCGATTTGGGCGAGCACCGTGGCTCATGCTGTAATTCAGCACTTT 773
DB CCCTTTAAATATGCGATTTGGGCGAGCACCTGTTGGCTCATGCTGTAATTTCCAGCACTTT 78378
QY 774 GGGAGGCGAGGTTGGTGAATCACTTGAGCCAGGAGTTTGACACGAGCTGCCCAACAT 893
DB GGGAGGCGAGGTTGGTGAATCACTTGAGCCAGGAGTTTGAGACGAGCTGCCCAACAT 78318
QY 834 GGTGAACACTCTGTCTTTACTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTCTGCT 893

78317 CGTGAACACTCTGTCTTTACT-----AAAAAAAAAAAAAAAAAAAAATCAGCCGGTGTGCG 78264
QY 894 GGGGCACACCTGTATATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTTGAACCTT 953
DB 78263 GGGGCACACCTGTATATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTTGAACCTT 78204
QY 954 GGAGGCGGAGTTTGCAAGTGGGCGGAGATCACATCACCGCCCTCCAGCTTGGGGCGACAGAG 1013
DB 78203 GGAGGTTGGAGTTTGCAAGTGGGCGGAGATCACATCACCGCCCTCCAGCTTGGGGCGACAGAG 78144
QY 1014 CAAGACTCTGTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1073
DB 78143 CAAGACTCTGTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 78084
QY 1074 TATATCT--AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1131
DB 78083 TATACCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 78024
QY 1132 AATGGGCCAAGAACTCTAAGGTATATTTGACAAATCATTTTCAAGAACCTTTAAAAAAGAAAG 1191
DB 78023 AATGGGCCAAGAACTCTAAGGTATATTTGACAAATCATTTTCAAGAACCTTTAAAAAAGAAAG 77964
QY 1192 AATCAGAGGCGATAGAAACAGAGGAGGAAACAGGGAGACAGAAACACCTTGTGCCCAAG 1251
DB 77963 AATCAGAGGCGATAGAAACAGAGGAGGAAACAGGGAGACAGAAACACCTTGTGCCCAAG 77904
QY 1252 GAGAACAAACAAAGGCTCTTAAGACAGACAGAG---GAGAGAGAGAGAGTGTAGTGA 1307
DB 77903 GAGAACAAACAAAGGCTCTTAAGACAGACAGAGGAGGCGGAGAGAGAGAGTGTAGTGA 77844
QY 1308 GAGACAGACAG 1347
DB 77843 GAGACAGACAG 77784
QY 1348 -----CAGAGAGACAGAGAGGCGGAGAGGAGATAGAAAGAGAGAGAGGAGGTTGA 1394
DB 77783 AGAGAGAGGCGGTGGAGAGAGACAGAGAGGCGGAGAGGAGATAGAAAGAGAGAGGAGGTTGA 77724
QY 1395 GAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGGAGAGACCAAGAGAGAG 1454
DB 77723 GAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGGAGAGACCAAGAGAGAG 77664
QY 1455 TCGAAGAGACTCTGAG 1514
DB 77663 TCGAAGAGACTCTGAG 77604
QY 1515 ACAGACAGTGGTGGAAATGAGCAAAATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1574
DB 77603 ACAGACAGTGGTGGAAATGAGCAAAATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77544
QY 1575 ATAGTGACCCAGAGTTGGTGAAGCCAGATCCTTAAGGCTGGGGGAGAGAGAGAGAGAGAGAG 1634
DB 77543 ATAGTGACCCAGAGTTGGTGAAGCCAGATCCTTAAGGCTGGGGGAGAGAGAGAGAGAGAGAG 77484
QY 1635 CTGGCTGGCTTCGGAGAGCCCTCCCATTTCTCCGGGCGCAGGAGGATAGGAGTGCACAT 1694
DB 77483 CTGGCTGGCTTCGGAGAGCCCTCCCATTTCTCCGGGCGCAGGAGGATAGGAGTGCACAT 77424
QY 1695 TCCGGAATGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGGAGAGAGAGAGAGAGAGAGAGAG 1754
DB 77423 TCCGGAATGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGGAGAGAGAGAGAGAGAGAGAGAG 77364
QY 1755 TAAAGCCCGATAGGCACTCATTTGCCGGGAATGTGCCCGAGGAGAGAGAGAGAGAGAGAGAGAG 1814
DB 77363 TAAAGCCCGATAGGCACTCATTTGCCGGGAATGTGCCCGAGGAGAGAGAGAGAGAGAGAGAGAG 77304
QY 1815 AACTCAGGCGCGGTGCCCGAG 1874
DB 77303 AACTCAGGCGCGGTGCCCGAG 77244
QY 1875 GAAGTCTGCGGTGAGCTGGGAGCAAAATCCCCACCCCTTACTCTGGGGGAGAGAGAGAGAGAGAG 1934


```
QY 4095 AGTAGCTGGATTACAGGTGCGCGCCACACACGCGTGGCTAAATTTTGTGTCTTTAGTAGA 4154
Db 75028 AGTAGCTGGATTACAGGTGCGCGCCACACACGCTAAATTTTGTGTCTTTAGTAGA 74969
QY 4155 GACAGGGTTTACCTTGTGGCCAGCGTGTCTCAAACTCCCAACTCCAGGCTCAGGTGATCGGCC 4214
Db 74968 GACAGGGTTTACCTTGTGGCCAGCGTGTCTCGAACTCCCAACTCCAGGCTCAGGTGATCGGCC 74909
QY 4215 CACCTCGGCCCTCCAGAGTGTCTGGGGTTACAGCGGTGAGCCACCGGCCCCAGGCCAAAGT 4274
Db 74908 CACCTCGGCCCTCCAGAGTGTCTGAGGTTACAGCGGTGAGCCACTGCCCCAGGCCAAAGT 74849
QY 4275 CAGAGCTCTTTATAGAGAGACTTAAATGTAACCTTGACCTGGCCCTTAACCTAAGTCAAT 4334
Db 74848 CAGAGCTC--TATAGAGAGACTTAAATGTAACCTTGACCTGGCCCTTAACCTAAGTCAAT 74791
QY 4335 TCCAAACCCCTTCCCTCGCTCCAGCCTGACCCCACTCACTGAGGCGTGCACCCACTTCTT 4394
Db 74790 TCCAAACCCCTTCCCTCGCTCCAGCCTGACCCCACTCACTGAGGCGTGCACCCACTTCTT 74731
QY 4395 GAGACCAAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCGAGGCTCCAGCCGCCAC 4454
Db 74730 GAGACCAAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCGAGGCTCCAGCCGCCAC 74671
QY 4455 AGCTTTGGCACTACCCCTGAGCTGTGTCAGGAATCTGTACCCAAATTTTACCCCTCAATG 4514
Db 74670 AGCTTTGGCACTACCCCTGAGCTGTGTCAGGAATCTGTACCCAAATTTTACCCCTCAATG 74611
QY 4515 TAGTTCTAGCCAAATTCAGGAATCTGTGAGGTCAGTTAGAGTCCAGTAAACCTTACTGA 4574
Db 74610 TAGTTCTAGCCAAATTCAGGAATCTGTGAGGTCAGTTAGAGTCCAGTAAACCTTACTGA 74551
QY 4575 GCCTGGGCTCTGTCTTGTAGCTTGAGCTGGGCTTGAGAGGTGCCACTCTTATTTCCAG 4634
Db 74550 GCCTGGGCTCTGTCTTGTAGCTTGAGGCTGGGCTTGAGAGGTGCCACTCTTATTTCCAG 74491
QY 4635 GCCTTGCCCTGCCCCCTCAGCATGTGACACCCACCCCTCTAGCTGGTGTGGCCTCTTG 4694
Db 74490 GCCTTGCCCTGCCCCCTCAGCGTGTGACACCCACCCCTCTAGCTGGTGTGGCCTCTTG 74431
QY 4695 AGTCTGAACCCAGCCCGCAGCCCAAGCCCGCTCTGAGCCCGCCGCGCAACCCATTTCCG 4754
Db 74430 AGTCTGAACCCAGCCCGCAGCCCAAGCCCGCTCTGAGCCCGCCGCGCAACCCATTTCCG 74371
QY 4755 TTCCAGAGCATCTTCTCGCAACAATGATGTTTCTGTGACCAACCCCTCTTAACACCGTG 4814
Db 74370 TTCCAGAGCATGTTTCTCGCAACAATGATGTTTCTGTGACCAACCCCTCTTAACACCGTG 74311
QY 4815 CCCTCTGGGAGCAACCAAGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGGTGCGATGAC 4874
Db 74310 CCCTCTGGGAGCAACCAAGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGGTGCGACGAC 74251
QY 4875 AGCAGAGCGCGATCATCATGATCGACTGCGATGATGACACCCAGCGGTGGAGGCC 4934
Db 74250 AGCAGAGCGCGATCATCATGATCGACTGCGATGATGACACCCAGCGGTGGAGGCC 74191
QY 4935 GCCTGTGTCTAAGGCGCCCAACAGCTCTACTCGGGCGCGTGTGTGTCATCCACAGTGG 4994
Db 74190 GCCTGTGTCTAAGGCGCCCAACAGCTCTACTCGGGCGCGTGTGTGTCATCCACAGTGG 74131
QY 4995 CTGCTC 5000
Db 74130 CTGCTC 74125
```

```
RESULT 9
AC130188/c
LOCUS
DEFINITION Papio anubis clone RP41-421P3, WORKING DRAFT SEQUENCE, 12 ordered
pieces.
ACCESSION AC130188
VERSION AC130188.2 GI:24960890
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
```

```
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

```
Papio anubis (olive baboon)
Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
1 (bases 1 to 176647)
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Carlaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-O., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stattipop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 176647)
Green,E.D.
Direct Submission
Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 176647)
Green,E.D.
Direct Submission
Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Nov 14, 2002 this sequence version replaced gi:22138439.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dsr
Center clone name: 421P03
```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171127 bases at least Q40
Consensus quality: 173424 bases at least Q30
Consensus quality: 174803 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 175547; sum-of-contigs
Quality coverage: 8.76x in Q20 bases; agarose-fp
Quality coverage: 9.39x in Q20 bases; sum-of-contigs

- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 12 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.
- * 1 6526: contig of 6526 bp in length
- * 6527 6626: gap of unknown length

*	6627	18812:	contig of 12186 bp in length
*	18613	18912:	gap of unknown length
*	18913	18913:	contig of 42868 bp in length
*	61781	61880:	gap of unknown length
*	61881	61880:	contig of 35188 bp in length
*	97069	97168:	gap of unknown length
*	97169	125206:	contig of 28038 bp in length
*	125307	125306:	gap of unknown length
*	125307	127021:	contig of 1715 bp in length
*	127022	129387:	gap of unknown length
*	129388	129487:	gap of unknown length
*	129488	132402:	contig of 2815 bp in length
*	132303	132402:	gap of unknown length
*	132403	135735:	contig of 3333 bp in length
*	135736	135835:	gap of unknown length
*	135836	143558:	contig of 7723 bp in length
*	143559	170182:	gap of unknown length
*	170183	170282:	gap of unknown length
*	170283	176647:	contig of 6365 bp in length.

[illegible]

Db	226	GCCAGAGCCAGGAGGACGGTGGCGCAGAAGGTGAGGCCTGGAATAATCCGTGGCTG	167
QY	1889	AGCTGGGAGCAAAATCCCCCCCACCTACCTGGGGGACAAGGCAACTGAGACCTGGTGAGG	1948
Db	166	AATGGNAGCATCCCCCACCCCTACCTACCTGGGGGACAAGGCAAGTGGGAGG	107
QY	1949	GTGGCTCAGCAGGAGGAGGAGAGGTGTGTGCCCTCTGCACCCACCATCTTTCTCTG	2008
Db	106	GTGGCTCAGCAGGAGGAGGAGGTGTGTGCATCTCTGCACCCACCATCTTTCTCTG	47
QY	2009	TC--CCCTCTTGGCCCTGTCTGAGGCTGTAGACTCTTAATTCTTCT	2052
Db	46	TCTTCCCTCTCGCCCGTCTGGAGGTGCTAGACTCTTAATTCTTCT	1
RESULT 10			
AC137591/c			
LOCUS	AC137591	32799 bp	DNA linear PRI 05-DEC-2002
DEFINITION	Homo sapiens chromosome X clone XX-B6cos map Xp22-PAR, complete sequence.		
ACCESSION	AC137591		
VERSION	AC137591.1	GI:25446705	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 32799)	Rump,A., Rosenthal,A., Drescher,B. and Schattevoy,R.	
AUTHORS	Direct Submission		
TITLE	Submitted (07-MAY-1997) Genome Analysis, Institute Of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany		
JOURNAL	2 (bases 1 to 32799)	Lagemann,D. and Platzter,M.	
REFERENCE	Direct Submission		
AUTHORS	Submitted (26-NOV-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany		
TITLE	3 (bases 1 to 32799)	Lagemann,D. and Platzter,M.	
JOURNAL	Direct Submission		
REFERENCE	Submitted (05-DEC-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany		
AUTHORS	----- Genome Center		
TITLE	Center: Insitute of Molecular Biotechnology		
JOURNAL	Center code: IMB		
COMMENT	Web site: http://genome.imb-jena.de/ Contact: gscj-submit@genome.imb-jena.de ----- Project Information Center project name: B6 Center clone name: XX-540F24 ----- Summary Statistics Sequencing vector: pUC18; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.90329 Consensus quality: 32765 bases at least Q40 Consensus quality: 32770 bases at least Q30 Consensus quality: 32770 bases at least Q20 Quality coverage:i3.12x ----- This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one pUC18 subclone. ----- Sequence Quality Assessment: This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.		

```

Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
FEATURES             Location/Qualifiers
     source            1..32799
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="X"
                        /map="Xp22-PAR"
                        /clone="XX-B6cos"
                        /c1one=".642"
     misc_feature      537..642
                        /note="single stranded/single chemistry region"
     misc_feature      2807..2820
                        /note="single stranded/single chemistry region"
     misc_feature      2913..2927
                        /note="low quality region"
     misc_feature      3166..3173
                        /note="single stranded/single chemistry region"
     misc_feature      5131..5234
                        /note="single stranded/single chemistry region"
     misc_feature      5235..5317
                        /note="single clone coverage"
     misc_feature      5266..5317
                        /note="low quality region"
     unsure            5266..5317
     misc_feature      5366..5431
                        /note="single clone coverage"
     misc_feature      5432..5518
                        /note="single stranded/single chemistry region"
     misc_feature      6151..6395
                        /note="single stranded/single chemistry region"
     unsure            6303
     unsure            6382
     misc_feature      7106..7137
                        /note="single stranded/single chemistry region"
     misc_feature      9265..9528
                        /note="single stranded/single chemistry region"
     variation         9669
                        /note="T substituted in clone: XX-P99cos"
                        /replace="c"
     variation         10317
                        /note="deleted in clone: XX-540F24"
                        /replace="n"
     variation         10897
                        /note="G substituted in clone: XX-1Bcos"
                        /replace="t"
     variation         11087..11090
                        /note="deleted in clone: XX-540F24"
                        /replace="n"
     variation         11259
                        /note="T substituted in clone: XX-540F24"
                        /replace="c"
     misc_feature      11744..11756
                        /note="low quality region"
     misc_feature      11765..11880
                        /note="single stranded/single chemistry region"
     variation         11834
                        /note="T substituted in clone: XX-540F24"
                        /replace="c"
     variation         11892..11900
                        /note="deleted in clone: XX-1Bcos"
                        /replace="n"
     variation         12197..12200
                        /note="deleted in clone: XX-1Bcos"
                        /replace="n"
     variation         12256..12259
                        /note="deleted in clone: XX-1Bcos"
                        /replace="n"
     variation         12277
                        /note="A substituted in clone: XX-540F24"
                        /replace="c"

```


[illegible]

[illegible]

RESULT 12

```

AC109997      141923 bp      DNA      linear      HTG 09-FEB-2002
LOCUS      Homo sapiens chromosome 5 clone CTD-207601, WORKING DRAFT SEQUENCE.
DEFINITION      3 unordered pieces.
ACCESSION      AC109997
VERSION      AC109997.1 GI:18642721
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 141923)
AUTHORS      DOE Joint Genome Institute.
TITLE      Sequencing of Human Chromosome 5
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 141923)
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Submitted (09-FEB-2002) Production Sequencing Facility, DOE Joint
COMMENT      Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 657070
Center clone name: CITB-H1_207601
-----
Summary Statistics
Consensus quality: 139579 bases at least Q40
Consensus quality: 140059 bases at least Q30
Consensus quality: 140264 bases at least Q20
Estimated insert size: 140000; agarose-fp estimation
Estimated insert size: 141723; sum-of-contigs estimation
Quality coverage: 8.57 in Q20 bases; agarose-fp estimation
Quality coverage: 8.47 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1097: contig of 1097 bp in length
* 1098 1197: gap of unknown length
* 1198 18670: contig of 17473 bp in length
* 18671 18770: gap of unknown length
* 18771 141923: contig of 123153 bp in length.
FEATURES
Location/Qualifiers
1..141923
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-207601"
/clone_lib="CalTech human BAC library D"
ORIGIN
Query Match 9.1%; Score 456.6; DB 2; Length 141923;
Best Local Similarity 79.0%; Pred. No. 1.6e-108;
Matches 573; Conservative 0; Mismatches 139; Indels 13; Gaps 2;
QY 326 GGTGGTCTATGATCATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACGGAGT 385
Db 91006 GATGGTTCATTTCTCAGGTTTTTTTTTTTGTGTTTTTTTTTTTGTGTTTTTTGAGACGGAGT 91065
QY 386 CTCGCTCTGTGCGCCAGGCTGAGTGACGTGCGGGATCTCGGCTCAGTCAAGCTCCGC 445
Db 91066 CTTGCTCTGTGGCCAGGTGGGAGTGCAGTGGCGCATCTCGCTCAGTCAAGCTCCGC 91125
QY 446 CTCGCCGGTTCACGCCCATCTCTCTCGCTCAGGCTCCCAAGTAGCTGGGACTACAGCGCC 505

```


2222	2308	7571	7564	13999	13754
-----	-----	-----	-----	-----	-----
360	<800	4492	4508	9092	9112
-----	-----	-----	-----	-----	-----
2344	2308	1092	1054	2356	2380
-----	-----	-----	-----	-----	-----
1198	1180	3332	3396	2287	2380
-----	-----	-----	-----	-----	-----
3828	3831	1123	1054	2669	2713
-----	-----	-----	-----	-----	-----
3328	3300	1803	1824	924	930
-----	-----	-----	-----	-----	-----
3248	3300	3041	3122	4641	4650
-----	-----	-----	-----	-----	-----
1289	1252	16514	16912	1247	1206
-----	-----	-----	-----	-----	-----
399	<800	192	<800	1708	1662
-----	-----	-----	-----	-----	-----
4134	4173	4763	4687	6082	6133
-----	-----	-----	-----	-----	-----
900	895	587	<800	1904	1897
-----	-----	-----	-----	-----	-----
3852	3963	5086	5041	4433	4394
-----	-----	-----	-----	-----	-----
4463	4378	2278	2421	6851	6941
-----	-----	-----	-----	-----	-----
3959	3831	1336	1314	1799	1783
-----	-----	-----	-----	-----	-----
1990	2005	3223	3122	23885	24086
-----	-----	-----	-----	-----	-----
4272	4173	7004	7057		
-----	-----	-----	-----		
2327	2308	971	1054		
-----	-----	-----	-----		
6858	6866	2341	2421		
-----	-----	-----	-----		
3056	3063	1649	1646		
-----	-----	-----	-----		
6815	6866	2822	2867		
-----	-----	-----	-----		
809	810	6084	6048		
-----	-----	-----	-----		
4433	4378	163	<800		
-----	-----	-----	-----		
6135	6032	9648	9389		
-----	-----	-----	-----		
		311	<800		
		3279	3396		
		3900	3950		
		2035	2029		
		3612	3630		
		3053	3229		
		3414	3396		
		865	883		
		637	<800		
		259	<800		
		97	<800		
		1242	1314		
		270	<800		
		23	<800		

Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.

Direct Submission

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (23-OCT-2001) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (15-JUN-2002) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

On Jun 15, 2002 this sequence version replaced gi:16328287.

COMMENT

Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: uwgchgs@u.washington.edu

----- Project Information

Center project name: Chr-3

Center clone name: RP11-349E16 (bc0646)

----- Summary Statistics

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator ET; 76% of reads

Chemistry: Dye-terminator Big Dye; 24% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 227062 bases at least Q40

Consensus quality: 227131 bases at least Q30

Consensus quality: 227137 bases at least Q20

Insert size: 227137; sum-of-contigs

Quality coverage: 8.0x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-728B21 (UWGC:bc0540) AC114876, 123790-bp overlap

3': RP11-286L5 (UWGC:bc0344) AC103559

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

NsII

EcoRI

HindIII

SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

22527	22843	8696	8633	3211	3196
7364	7368	6	<800	6382	6483
1250	1186	4641	4671	512	<800
1646	1681	924	937	449	<800
5862	5809	2669	2702	1398	1327
4415	4376	2286	2141	1649	1684
5203	5251	2356	2346	2341	2357
2339	2380	9092	9731	971	998
4104	4145	13999	13681	7003	7067
2067	2088	6660	6723	3223	3196
1413	1416	286	<800	1336	1327
8832	8730	8529	8633	2278	2251
1451	1416	11874	11716	5086	5023
3108	3003	10800	10477	587	<800
1155	1088	2842	2850	4763	4697
10132	9934	5654	5704	192	<800
7178	7368	8640	8633	16514	16792
2392	2380	13221	12924	3041	3013
4819	4786	555	<800	1803	1813
1188	1186	5455	5489	1123	1071
2666	2689	3694	3751	3332	3382
1188	1186	4919	4887	1092	1071
431	<800	951	937	4492	4528
732	750	334	<800	7571	7668
1680	1681	3403	3506	4767	4697
1855	1877	3587	3633	817	832
1537	1490	2819	2850	1340	1327
1361	1350	3838	3909	1004	998
5212	5251	5684	5704	891	922
24	<800	373	<800	10593	10357
1608	1681	63	<800	2970	2846
1731	1877	203	<800	9418	9439
1439	1416	7896	7870	1527	1526
327	<800	332	<800	2795	2775
5312	5251	407	<800	1052	1071
3221	3262	3904	3909	305	<800

Db	91802	GCTCATCTAATGAAGAAATTATGACAGGGTGTAAAACTAGGCAGGGGATGGGCATCA	91861
Qy	746	TGCTCATGCTCTAATCCAGCACATTTGGGAGGCAGAGTGGGTGAATCACTCTCAGGCC	805
Db	91862	TGGCTCAGCCTGTAAATCCAGCACATTTGGGAGGCCAAGCCAGTGGATCACTCAGGTC	91921
Qy	806	AGGAGTTTGAGACCGCTGGCCAACATGGTCGAACTCTGCTTTACTAAAAAATAAAAA	865
Db	91922	CGAGTTTCGAGACCATCTGGCCACGTCGAAACCCCATCTCTACT	91969
Qy	866	AAAAAAAAAAAAATCAGCCGGGTGTCTGGGGCACACCTGTAAATCCAGCTATGCTGGAG	925
Db	91970	AAAAATACAAAATTAGCAGGCTTAGTGGTGTGTGCTGTGTAGTCCAGCTACTCGGNG	92029
Qy	926	GCTGAGCAGAGAGTCATGTGAACCTCGAGGCGGAGGTTGCAGTGGGCCGAGATCACA	985
Db	92030	TC TGAGGCATGAGAATCACTTGAACCTGGAGGTTCGAGTTCGAGCCCAAGATTACG	92089
Qy	986	TCACCGCCCTCCAGCCTGGCGCAGAGCAGACGACTCTGTCTCAAAATAAATAATAACAA	1045
Db	92090	CCACTGTACTCCAGCCTGGGTGACAGAGCAAGACTCTGTCTCAAAAAATAAATAATAA	92149
Qy	1046	ACGAA 1050	
Db	92150	AATAA 92154	

Search completed: July 4, 2005, 09:55:23
Job time : 14430.2 secs

[illegible]

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 1759.59 Seconds

(without alignments)
16821.327 Million cell updates/sec

Title: US-09-936-271c-13_COPY_1_5000

Perfect score: 5000

Sequence: 1 gggccagagtggaagcaag,.....tgatccacacagtggctgc 5000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_16Dec04:*

- 1: Geneseq_1980s:*
- 2: Geneseq_1980s:*
- 3: Geneseq_2000s:*
- 4: Geneseq_2001as:*
- 5: Geneseq_2001bs:*
- 6: Geneseq_2002as:*
- 7: Geneseq_2002bs:*
- 8: Geneseq_2003as:*
- 9: Geneseq_2003bs:*
- 10: Geneseq_2003cs:*
- 11: Geneseq_2003ds:*
- 12: Geneseq_2004as:*
- 13: Geneseq_2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5000	100.0	11570	3 AAA95905	AAA95905 Human KLK
2	5000	100.0	11570	12 ADK52482	Adk52482 Human kal
3	5000	100.0	11570	13 ADR72623	Adr72623 Human ren
4	5000	100.0	11570	13 ADR72875	Adr72875 Human ova
5	472	9.4	525	12 ACH80345	Ach80345 Human gen
6	440.2	8.8	41454	11 ACN44942	Acn44942 Human gen
7	424.4	8.5	84607	2 AAX90847	Aax90847 Human PAC
8	419	8.4	91352	12 ADN94799	Adn94799 DNA encod
9	418.4	8.4	115935	13 ADR52890	Adr52890 Drug ther
10	417.6	8.4	26371	11 ACN44814	Acn44814 Human gen
11	417.6	8.4	91760	11 ACN44410	Acn44410 Human gen
12	417.4	8.3	216215	10 ADF69167	Adf69167 Human MP5
13	416.6	8.3	139904	6 ABK83562	Abk83562 Human cDN
14	416.2	8.3	79652	10 ADG70184_3	Continuation (4 of
15	415.2	8.3	17397	4 AAS36445	Aas36445 Human car
16	415.2	8.3	17397	10 ADE47139	Ade47139 Human car
17	415.2	8.3	17397	13 ADJ08557	Adj08557 Human car
18	415.2	8.3	19334	4 AAS36443	Aas36443 Human car
19	415.2	8.3	19334	10 ADE47137	Ade47137 Human car
20	415.2	8.3	19334	13 ADJ08555	Adj08555 Human car

21	415.2	8.3	19345	4 AAS36444	Aas36444 Human car
22	415.2	8.3	19345	10 ADE47138	Ade47138 Human car
23	415.2	8.3	19345	13 ADJ08556	Adj08556 Human car
C 24	415.2	8.3	29001	12 ADO85575	Ado85575 Human zin
C 25	414.8	8.3	12970	4 AAL05001	Aal05001 Human rep
C 26	414.8	8.3	12970	4 ABL97894	Ab197894 Human tes
C 27	414.6	8.3	74037	6 ABK94412	Abk94412 DNA encod
C 28	414.6	8.3	111084	12 ADQ18808	Adq18808 Human sof
29	414.2	8.3	212231	11 ACM44598	Acn44598 Human gen
30	410.4	8.2	10445	4 AAL03692	Aal03692 Human rep
C 31	409.6	8.2	50000	9 ADB16926	Adb16926 Human DYX
C 32	408.2	8.2	240000	8 ACD13446	Acd13446 Human DNA
C 33	407.6	8.2	196686	11 ACN44170	Acn44170 Human gen
C 34	407.6	8.1	32460	12 ADQ07738	Adq07738 Nucleotid
35	406.2	8.1	12970	4 AAL05001	Aal05001 Human rep
36	406.2	8.1	12970	4 ABL97894	Ab197894 Human tes
C 37	406	8.1	32224	4 AAK89986	Aak89986 Human dig
C 38	405.8	8.1	4433	4 AAL04142	Aal04142 Human rep
C 39	405.8	8.1	4433	5 ABA18400	Aba18400 Human ner
C 40	405.8	8.1	4433	5 AAS40516	Aas40516 DNA encod
C 41	405.8	8.1	4433	11 ADJ09722	Adj09722 Human pro
42	405.8	8.1	126001	12 ADH77123	Adh77123 Human PAZ
43	405	8.1	32767	8 AAD56075	Aad56075 Human SNL
44	405	8.1	32767	9 ADA02437	Ada02437 Human SNL
45	405	8.1	32767	10 ADB72176	Adb72176 Human SNL

ALIGNMENTS

RESULT 1
AAA95905
ID AAA95905 standard; DNA; 11570 BP.
XX
AC AAA95905;
XX
DT 02-FEB-2001 (first entry)
XX
DE Human KLK-L2 gene.
XX
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KW kallikrein-like protein; serine protease; cytosolic; cancer;
KW prostrate cancer; ds.
XX
OS Homo sapiens.
XX
XX WO2000053776-A2.
XX
PD 14-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-CA000258.
XX
PR 11-MAR-1999; 99US-0124260P.
PR 01-APR-1999; 99US-0127386P.
PR 21-JUL-1999; 99US-0144919P.
XX
PA (MOUN) MOUNT SINAI HOSPITAL.
XX
PI Yousef GM, Diamandis EP;
XX
DR WPI; 2000-587440/55.
DR P-PSDB; AAB21296.
XX
PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
XX protein mediated disorders, especially cancer.
XX
PS Claim 1; Page 143-149; 184pp; English.
XX
CC The present sequence is the coding sequence of the human KLK-L2 gene,
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like
CC proteins are a subgroup of the serine protease enzyme family. They
CC catalyse the selective cleavage of specific polypeptide precursors to
CC release peptides with potent biological activity. Nucleic acids encoding

Db 1981 GTGGGTCTGACCAACATCTTTCTGTGTCCTTCCCTTCTGAGGCTGCTAGA 2040
Qy 2041 CTCCTATCTTCTGAATTCATAGTCCCTGCTGCTCAGCGAGTCCCATGCTGCCCGTC 2100
Db 2041 CTCCTATCTTCTGAATTCATAGTCCCTGCTGCTCAGCGAGTCCCATGCTGCCCGTC 2100
Qy 2101 CTTGTGTTCTCTCTACCTGGGGAATAAGTAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db 2101 CTTGTGTTCTCTCTACCTGGGGAATAAGTAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Qy 2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGTGCAGCGGCC 2220
Db 2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGTGCAGCGGCC 2220
Qy 2221 ATGGCTACAGCAAGACCCCCCTGATGTTGGTCTCTGTCTGTCTGTATCAGAGCTTGT 2280
Db 2221 ATGGCTACAGCAAGACCCCCCTGATGTTGGTCTCTGTCTGTCTGTATCAGAGCTTGT 2280
Qy 2281 CTGGGGTCAAGAGTACCAAGACTCTGGGGTGGGAGGTTGTGGGATTTGGAGGACTGT 2340
Db 2281 CTGGGGTCAAGAGTACCAAGACTCTGGGGTGGGAGGTTGTGGGATTTGGAGGACTGT 2340
Qy 2341 CTCTGGGCACTAGAGCGCTGTCCCTGGGAACTGTGTGAGCTGGGATGACTCCGG 2400
Db 2341 CTCTGGGCACTAGAGCGCTGTCCCTGGGAACTGTGTGAGCTGGGATGACTCCGG 2400
Qy 2401 GACCGGTGAATGAGTCTCTGTCTGTACTTGTGTTGTCGATCGTATGTGCCCTGT 2460
Db 2401 GACCGGTGAATGAGTCTCTGTCTGTACTTGTGTTGTCGATCGTATGTGCCCTGT 2460
Qy 2461 GACTGCCACGTTGTGTGGGAGGAGGATGCTCTTTCCATATCAGTGTGCGG 2520
Db 2461 GACTGCCACGTTGTGTGGGAGGAGGATGCTCTTTCCATATCAGTGTGCGG 2520
Qy 2521 CAGTGGCACTGACCTTTGAGGCTGTGTGTGCTGTTTGTGATTTGATTTAAAG 2580
Db 2521 CAGTGGCACTGACCTTTGAGGCTGTGTGTGCTGTTTGTGATTTGATTTAAAG 2580
Qy 2581 ATTGTGTGGCTCCACAGCTGTGTGGTGAATGATGATGATGATGATGATGATGAT 2640
Db 2581 ATTGTGTGGCTCCACAGCTGTGTGGTGAATGATGATGATGATGATGATGATGAT 2640
Qy 2641 GTGTTGGCTGTGTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
Db 2641 GTGTTGGCTGTGTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
Qy 2701 TCCCTGAGTCCCGGATTCGTCGCAACAAAGTGTGATCACCATGGAAGCTGTGACT 2760
Db 2701 TCCCTGAGTCCCGGATTCGTCGCAACAAAGTGTGATCACCATGGAAGCTGTGACT 2760
Qy 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTTGGCTGAGTGTGATGATGATGATGATGAT 2820
Db 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTTGGCTGAGTGTGATGATGATGATGATGAT 2820
Qy 2821 TTTGTGACCGGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTGTGTGTGTGTGT 2880
Db 2821 TTTGTGACCGGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTGTGTGTGTGTGT 2880
Qy 2881 TCTGTGTAGCGCGTGAATGCTATCTGTATGTGTGATGATGATGATGATGATGATGATGAT 2940
Db 2881 TCTGTGTAGCGCGTGAATGCTATCTGTATGTGTGATGATGATGATGATGATGATGATGAT 2940
Qy 2941 TTTCTGTCTCTCCCTGGAGGATAGAGGTTGAGGGTGTAGCTATCTCTGGGAGTGGGTGC 3000
Db 2941 TTTCTGTCTCTCCCTGGAGGATAGAGGTTGAGGGTGTAGCTATCTCTGGGAGTGGGTGC 3000
Qy 3001 CAGTGTGACTGACTGTGCTGTGTGTGCAAGAGTATGTGGCAGTCTGAAATC 3060
Db 3001 CAGTGTGACTGACTGTGCTGTGTGTGCAAGAGTATGTGGCAGTCTGAAATC 3060
Qy 3061 TGTGCACACGCGATCTGCTGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Db 3061 TGTGCACACGCGATCTGCTGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 3120

Db 3061 TGTGCACACGCGATCTGCTGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Qy 3121 GCTAGGCTGCTCCCGGAGCGTGTGATCCTGAGACAGAGCTGTATGTTAGCTGCACTGTGTG 3180
Db 3121 GCTAGGCTGCTCCCGGAGCGTGTGATCCTGAGACAGAGCTGTATGTTAGCTGCACTGTGTG 3180
Qy 3181 GAGCAACATGGCGTGTCTGCAAGAACTGCTGCTGCTTGGCTGTGTTACTGCTGCTTGTGC 3240
Db 3181 GAGCAACATGGCGTGTCTGCAAGAACTGCTGCTGCTTGGCTGTGTTACTGCTGCTTGTGC 3240
Qy 3241 GCGTGGTCTTGGGGTGAATTCGTAATGATGTTGGTGTGCTGAGGCGCATCAGCAAGGTA 3300
Db 3241 GCGTGGTCTTGGGGTGAATTCGTAATGATGTTGGTGTGCTGAGGCGCATCAGCAAGGTA 3300
Qy 3301 GAACAGGCGCGGCGGCTGCTCAAGCTGTGTAATCCAGCCCTTTGGGAGGCGGAGCA 3360
Db 3301 GAACAGGCGCGGCGGCTGCTCAAGCTGTGTAATCCAGCCCTTTGGGAGGCGGAGCA 3360
Qy 3361 GCGGATCAGCTGAGTTCGGGAGATCGAGGCGGCTGACCAACATGAGAACCCCTCT 3420
Db 3361 GCGGATCAGCTGAGTTCGGGAGATCGAGGCGGCTGACCAACATGAGAACCCCTCT 3420
Qy 3421 CTACTAAAAATACAAAAATTTAGCTGTGTGTGGCGCTGCTGTATATCCAGCTACTC 3480
Db 3421 CTACTAAAAATACAAAAATTTAGCTGTGTGTGGCGCTGCTGTATATCCAGCTACTC 3480
Qy 3481 GCGAGACTGGGCGAGAAAAATTCGTTGAAACCGGAGGTGGAGTTCCGCTGAGCCGAGA 3540
Db 3481 GCGAGACTGGGCGAGAAAAATTCGTTGAAACCGGAGGTGGAGTTCCGCTGAGCCGAGA 3540
Qy 3541 TCGGCGATTTGCACTCCAGCTGGGCAACAGAGCGGAACTCCGCTCTCGAAGAAAAA 3600
Db 3541 TCGGCGATTTGCACTCCAGCTGGGCAACAGAGCGGAACTCCGCTCTCGAAGAAAAA 3600
Qy 3601 GAAAAAAGGTGAAGAACTGTAATGGGCAACGAGGAGCTGATGATGAGTGGGCG 3660
Db 3601 GAAAAAAGGTGAAGAACTGTAATGGGCAACGAGGAGCTGATGATGAGTGGGCG 3660
Qy 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGATTTGACAGGATTTGAGAGGAT 3720
Db 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGATTTGACAGGATTTGAGAGGAT 3720
Qy 3721 GTTTTCACTGTAGATTTCAAAACCTAGGCTGTCTTCCCTCCATGTTGGCCCTTAAG 3780
Db 3721 GTTTTCACTGTAGATTTCAAAACCTAGGCTGTCTTCCCTCCATGTTGGCCCTTAAG 3780
Qy 3781 CTGAGCCCTTCTTCTGTGCTTTCGAAACCTAGCTCCGCGCATGAGCTCTGACC 3840
Db 3781 CTGAGCCCTTCTTCTGTGCTTTCGAAACCTAGCTCCGCGCATGAGCTCTGACC 3840
Qy 3841 CCACCTCTTCTTCTCAACACGCGCTTAGGCGAGCTCTAGTGGACCCCGCTAAGGCA 3900
Db 3841 CCACCTCTTCTTCTCAACACGCGCTTAGGCGAGCTCTAGTGGACCCCGCTAAGGCA 3900
Qy 3901 CACCCCTTTGGGCGAGCTCCACCCCTATTTGTGGGTACCTTCTAGAACCCCTTCAA 3960
Db 3901 CACCCCTTTGGGCGAGCTCCACCCCTATTTGTGGGTACCTTCTAGAACCCCTTCAA 3960
Qy 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGTCTCTCTCCAGGCTGGAG 4020
Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGTCTCTCTCTCCAGGCTGGAG 4020
Qy 4021 TGCAGTGGCTGTATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGTCTCGT 4080
Db 4021 TGCAGTGGCTGTATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGTCTCGT 4080
Qy 4081 GCCTTCAACCTCTCTAGTGTGGATTTACAGGTGCGCGCCACACGCTGGCTAATTTT 4140
Db 4081 GCCTTCAACCTCTCTAGTGTGGATTTACAGGTGCGCGCCACACGCTGGCTAATTTT 4140
Qy 4141 GTGTCTTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGTCTCAAACTCCCAACC 4200
Db 4141 GTGTCTTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGTCTCAAACTCCCAACC 4200

QY 4201 TCAGGTGATCGGCCACCTCGGCCTCCAGAGTCTGGGGTTACAGCGTGGAGCCACCGC 4260
Db TCAGGTGATCGGCCACCTCGGCCTCCAGAGTCTGGGGTTACAGCGTGGAGCCACCGC 4260
QY 4261 CCCAGGCCAAAGTCAGAGCTCTTTATAGAGAGCTCTAAATGTAACCCCTGAGCCCTGGCC 4320
Db CCCAGGCCAAAGTCAGAGCTCTTTATAGAGAGCTCTAAATGTAACCCCTGAGCCCTGGCC 4320
QY 4321 CTAACCTAAGTCAATTCCTCAAAACCCCTCGCTCCAGCCCTGAGCCCTGAGCCCTGAGCC 4380
Db CTAACCTAAGTCAATTCCTCAAAACCCCTCGCTCCAGCCCTGAGCCCTGAGCCCTGAGCC 4380
QY 4381 TGACCCCACTCTTTGAGACAGAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440
Db TGACCCCACTCTTTGAGACAGAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440
QY 4441 GCTCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAAT 4500
Db GCTCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAAT 4500
QY 4501 TTTACCCCTCAGATGATCTAGCAATTCAGGAATCTCTGAGGTCCAGTTAGATCCA 4560
Db TTTACCCCTCAGATGATCTAGCAATTCAGGAATCTCTGAGGTCCAGTTAGATCCA 4560
QY 4561 GTAAACCTTACCTGAGCCTGGGCTCTGTCTTGTAGCTTGGCTGGGCTTGAGAGGTGCCA 4620
Db GTAAACCTTACCTGAGCCTGGGCTCTGTCTTGTAGCTTGGCTGGGCTTGAGAGGTGCCA 4620
QY 4621 CTCTTATCTCAGGCTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
Db CTCTTATCTCAGGCTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
QY 4681 GGTCTGCTCTTGTAGTCTGAACCCACCCAGCCCGCCAGCCCGCCAGCCCGCCAGCCCGCC 4740
Db GGTCTGCTCTTGTAGTCTGAACCCACCCAGCCCGCCAGCCCGCCAGCCCGCCAGCCCGCC 4740
QY 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTGTGACCAAC 4800
Db CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTGTGACCAAC 4800
QY 4801 CCTCTAACACCGTCCCTCTGGAGCAACACGAGACCTGGAGCTGGGGCGGGGAGAGC 4860
Db CCTCTAACACCGTCCCTCTGGAGCAACACGAGACCTGGAGCTGGGGCGGGGAGAGC 4860
QY 4861 CCCGTCGATGACAGCAGCAGCGGCATCATCAATGATCCGACTGCGATATGCACACC 4920
Db CCCGTCGATGACAGCAGCAGCGGCATCATCAATGATCCGACTGCGATATGCACACC 4920
QY 4921 AGCGTGGCAGGCGCGCTGTTGTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTGG 4980
Db AGCGTGGCAGGCGCGCTGTTGTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTGG 4980
QY 4981 TGCATCCACAGTGGCTGCTC 5000
Db TGCATCCACAGTGGCTGCTC 5000

RESULT 2

ADK52482
ID ADK52482 standard; DNA; 11570 BP.

AC ADK52482;

XX 03-JUN-2004 (first entry)

DE Human kallikrein 5 encoding sequence.

XX kallikrein 5; cancer; Cytostatic; Immunostimulant; ds.

XX Homo sapiens.

OS WO2004021008-A2.

XX 11-MAR-2004.
XX 28-AUG-2003; 2003WO-CA001310.
XX 28-AUG-2002; 2002US-0407333P.
XX (MOUN) MOUNT SINAI HOSPITAL.
XX (YOUS/) YOUSEF G.
XX Diamandis EP;
XX WPI; 2004-239232/22.
XX P-PSDB; ADK52481.
XX Detecting kallikrein 5 associated with cancer, useful in diagnosing,
XX monitoring, detecting, imaging and treating breast or ovarian carcinoma,
XX comprises comparing the detected amount of kallikrein 5 in a sample with
XX a standard sample.
XX Disclosure; SEQ ID NO 2; 59pp; English.
XX The present inventin relates to detecting kallikrein 5 associated with
XX breast or ovarian cancer in a patient comprises detecting in the sample
XX kallikrein 5 and comparing the detected amount with an amount detected
XX for a standard. The method is useful in detecting kallikrein 5 associated
XX with breast or ovarian cancer in a patient. The methods and kits are
XX useful in diagnosing, monitoring, detecting, imaging and treating breast
XX or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for
XX preventing and treating breast and ovarian cancer and for stimulating or
XX enhancing antibody production or for inducing an immune response. The
XX present sequence represents human kallikrein 5 encoding sequence.
XX SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 5000; DB 12; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCAGAGTGAAGCAAGAGAGAGTGTAGAGCTCCCTCTGCAAAAGTGGTTGAGTC 60
Db 1 GGGCCAGAGTGAAGCAAGAGAGAGTGTAGAGCTCCCTCTGCAAAAGTGGTTGAGTC 60
QY 61 TCCCTCTGCTTAAATCGAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 61 TCCCTCTGCTTAAATCGAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 AAGAAAG 180
Db 121 AAGAAAG 180
QY 181 ACAGAGAGCTGGGACACAGGGACACACAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGGACACACAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 ACACAATGGAGACACAGAGAGTGTAAAGAAAGAGAGATTAACAGAGTCCCAGATACACGC 300
Db 241 ACACAATGGAGACACAGAGAGTGTAAAGAAAGAGAGATTAACAGAGTCCCAGATACACGC 300
QY 301 AAAGGGCAGAGACACAGTGTTCAGGGTGGTGTCTATGATCATCTCTCTCTCTCTCTCT 360
Db 301 AAAGGGCAGAGACACAGTGTTCAGGGTGGTGTCTATGATCATCTCTCTCTCTCTCTCT 360
QY 361 TTTTCT 420
Db 361 TTTTCT 420
QY 421 GATCTCGGCTCACTGCAAGCTCGGCTCCGGGTTCAAGCCATTCCTCTCTCTCTCTCTCT 480
Db 421 GATCTCGGCTCACTGCAAGCTCGGCTCCGGGTTCAAGCCATTCCTCTCTCTCTCTCTCT 480
QY 481 CCAAGTAGCTGGAGCTACAGGGCCCGCCACTACGCCCGGCTAAATTTTGTATTTTAA 540
Db 481 CCAAGTAGCTGGAGCTACAGGGCCCGCCACTACGCCCGGCTAAATTTTGTATTTTAA 540

Db 481 CCAAGTAGCTGGGACTACAGCGCGCCGCCACTACGCCCGGCTAAATTTTTTTGTAATTTTA 540
Qy 541 GTAGAGAGGGGTTTCAACCGTTTATAGCCGGGATGGGCTCGATCTCTGACCTCGTATCC 600
Db 541 GTAGAGAGGGGTTTCAACCGTTTATAGCCGGGATGGGCTCGATCTCTGACCTCGTATCC 600
Qy 601 GCCCGCTCGGCTCCCAAGTGTGGGATTTACAGGCGTGGAGCCACCGCCGCCGCCCATG 660
Db 601 GCCCGCTCGGCTCCCAAGTGTGGGATTTACAGGCGTGGAGCCACCGCCGCCGCCCATG 660
Qy 661 ATCATCTTCTTGACTATGCTGATGTGACAAAGTACTTAAAGCCATCAGACTCTACCCCTTTA 720
Db 661 ATCATCTTCTTGACTATGCTGATGTGACAAAGTACTTAAAGCCATCAGACTCTACCCCTTTA 720
Qy 721 AATATGACAGTTTGGGCCAGGACCGTGGCTCATGCCCTGTAAATTCACGACCTTTGGGAGGC 780
Db 721 AATATGACAGTTTGGGCCAGGACCGTGGCTCATGCCCTGTAAATTCACGACCTTTGGGAGGC 780
Qy 781 AGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGCCTGGCCCAACATGGTGAAA 840
Db 781 AGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGCCTGGCCCAACATGGTGAAA 840
Qy 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900
Db 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900
Qy 901 ACCTGTAATCCAGTATGCTGGAGGCTGAGGACGAGAGTCACTTGAACCTTGAGGCG 960
Db 901 ACCTGTAATCCAGTATGCTGGAGGCTGAGGACGAGAGTCACTTGAACCTTGAGGCG 960
Qy 961 GAGGTTGAGTGGCGGAGATCAATCACCCTCCAGGCTGGGCGACAGAGCAAGACT 1020
Db 961 GAGGTTGAGTGGCGGAGATCAATCACCCTCCAGGCTGGGCGACAGAGCAAGACT 1020
Qy 1021 CTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080
Db 1021 CTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080
Qy 1081 AAAAAAATAATGCTGTCAACAAATAGAGCAAGTGAATAAAGGAAATAAATAATGGGCCA 1140
Db 1081 AAAAAAATAATGCTGTCAACAAATAGAGCAAGTGAATAAAGGAAATAAATAATGGGCCA 1140
Qy 1141 AGAATCTAAGGTATATTTGCAAAATCATTTCAGAACTTTTAAAGAAAGAAATACAGAG 1200
Db 1141 AGAATCTAAGGTATATTTGCAAAATCATTTCAGAACTTTTAAAGAAAGAAATACAGAG 1200
Qy 1201 GGCAATAGAAACAGAGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAGAAACAA 1260
Db 1201 GGCAATAGAAACAGAGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAGAAACAA 1260
Qy 1261 ACAAGGCTCTTAAGACAGACAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 ACAAGGCTCTTAAGACAGACAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 GAGAGAGGGTGGAGAGAGACACAGATATTTGAGAGAGACTCAGAAAGATAGCCGAGGGA 1440
Db 1381 GAGAGAGGGTGGAGAGAGACACAGATATTTGAGAGAGACTCAGAAAGATAGCCGAGGGA 1440
Qy 1441 GAACCCACAGAGATGGAAGAGAGACTCTGAGAAATAAACCAGAGACAAAGATGGAAGAGG 1500
Db 1441 GAACCCACAGAGATGGAAGAGAGACTCTGAGAAATAAACCAGAGACAAAGATGGAAGAGG 1500
Qy 1501 AGTATCAGAGGTGAACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560
Db 1501 AGTATCAGAGGTGAACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560
Qy 1561 TCCAGGCGCCAGAAATAGTAGTACCCAGAGTGGTGAAGCCAGATCCTTAAGGCTGGGG 1620
Db 1561 TCCAGGCGCCAGAAATAGTAGTACCCAGAGTGGTGAAGCCAGATCCTTAAGGCTGGGG 1620

Qy 1621 AGCAGGGAGGGGCTGGCTGGCTTCCGAGAGACCCCTCCCATTTCTCCGGCCAGGGAG 1680
Db 1621 AGCAGGGAGGGGCTGGCTGGCTTCCGAGAGACCCCTCCCATTTCTCCGGCCAGGGAG 1680
Qy 1681 GTAGGAGAGTACATTCCTGGACTGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGCA 1740
Db 1681 GTAGGAGAGTACATTCCTGGACTGGGTGGGGGTGCTCTGGGGGTGGAGATAGGGGGAGCA 1740
Qy 1741 GGAGGAGCTAATTCCTAAGCCCGATAGGCACTCATTTGCCCCGGGAATGTGCCCCAGGGAG 1800
Db 1741 GGAGGAGCTAATTCCTAAGCCCGATAGGCACTCATTTGCCCCGGGAATGTGCCCCAGGGAG 1800
Qy 1801 CAGTGGGTGGTATAACTCAGGCCCGGTGCCAGAGCCAGAGAGGAGGAGTGGCCAGGA 1860
Db 1801 CAGTGGGTGGTATAACTCAGGCCCGGTGCCAGAGCCAGAGAGGAGGAGTGGCCAGGA 1860
Qy 1861 AGGCACAGGGCTCAGAAAGTCTGGGGTCTGAGCTGGGAGCAAAATCCCCCACCCTTACCTGG 1920
Db 1861 AGGCACAGGGCTCAGAAAGTCTGGGGTCTGAGCTGGGAGCAAAATCCCCCACCCTTACCTGG 1920
Qy 1921 GGACACAGGGCAAGTGAGACCTGTGAGGGTGGCTCAGCAGCGAGGGAAGAGAGGTGTCT 1980
Db 1921 GGACACAGGGCAAGTGAGACCTGTGAGGGTGGCTCAGCAGCGAGGGAAGAGAGGTGTCT 1980
Qy 1981 GTGGTCTGACACCAATCTTCTGTCTCCCTCTGTGGCCCTGTCTGGAGGCTGTCTAGA 2040
Db 1981 GTGGTCTGACACCAATCTTCTGTCTCCCTCTGTGGCCCTGTCTGGAGGCTGTCTAGA 2040
Qy 2041 CTCTATCTTCTCAATTTATAGTGGTCTCAGCGAGTGGCCGATGTGGCCCGTCT 2100
Db 2041 CTCTATCTTCTCAATTTATAGTGGTCTCAGCGAGTGGCCGATGTGGCCCGTCT 2100
Qy 2101 CTGTGGTCTCTCTACCTGGGGAAATAAGTGTAGGGAGGAGGAGGAAAGTGGGTTAAGG 2160
Db 2101 CTGTGGTCTCTCTACCTGGGGAAATAAGTGTAGGGAGGAGGAGGAAAGTGGGTTAAGG 2160
Qy 2161 GCTCCCCGAGTCCGCTGGGCTCCCAACCTCTGACATTTCCCATTCAGGTGACGGGCC 2220
Db 2161 GCTCCCCGAGTCCGCTGGGCTCCCAACCTCTGACATTTCCCATTCAGGTGACGGGCC 2220
Qy 2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTGCTGTCTGTCTGTATCAGACGCTTCTCT 2280
Db 2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTGCTGTCTGTCTGTATCAGACGCTTCTCT 2280
Qy 2281 CTGGGGTCAACAGGTAAACAGAACTCTGGGGTGGAGGGTTGTGGGATTTGGAGGACTGT 2340
Db 2281 CTGGGGTCAACAGGTAAACAGAACTCTGGGGTGGAGGGTTGTGGGATTTGGAGGACTGT 2340
Qy 2341 CTCTGGGCACTAGAGCGCTGTCCCTGGGGAACTGTGTGAGCCTGGGCATGACTCCGG 2400
Db 2341 CTCTGGGCACTAGAGCGCTGTCCCTGGGGAACTGTGTGAGCCTGGGCATGACTCCGG 2400
Qy 2401 GACCGGTGAATGTGAGTCTCTGTCTGTCTGTGGTGTGGGATTCGATGTGGGCCCTGT 2460
Db 2401 GACCGGTGAATGTGAGTCTCTGTCTGTCTGTGGTGTGGGATTCGATGTGGGCCCTGT 2460
Qy 2461 GACTGCCACCGTGTGTCTGGGGAGGGGATGCTTTTCCCATATCAGAGTGTGTGCGG 2520
Db 2461 GACTGCCACCGTGTGTCTGGGGAGGGGATGCTTTTCCCATATCAGAGTGTGTGCGG 2520
Qy 2521 CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2521 CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Qy 2581 ATTGTGTGGCTCCACAGCTGTGTGGTGAATGATGATGATGATGATGATGATGATGATGATG 2640
Db 2581 ATTGTGTGGCTCCACAGCTGTGTGGTGAATGATGATGATGATGATGATGATGATGATGATG 2640
Qy 2641 GTGTTTGGCTGT 2700
Db 2641 GTGTTTGGCTGT 2700

QY 2701 TCCCTGAGGTCCCGGATTTGGTGCACAAAGTGTCTATCAACATGGAAGCTGTGACT 2760
DB 2701 TCCCTGAGGTCCCGGATTTGGTGCACAAAGTGTGTATCAACATGGAAGCTGTGACT 2760
QY 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTTGGCTGAGTGTGACGTTATGATGCCGTA 2820
DB 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTTGGCTGAGTGTGACGTTATGATGCCGTA 2820
QY 2821 TTTGTGACCGTGTGACTTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG 2880
DB 2821 TTTGTGACCGTGTGACTTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG 2880
QY 2881 TCTGTGTAGCGCGTGTAAATGTCTATGTATGTGTGATGTGAGCTGTGTGTGAGT 2940
DB 2881 TCTGTGTAGCGCGTGTAAATGTCTATGTATGTGTGATGTGAGCTGTGTGTGAGT 2940
QY 2941 TTCTGTCTCTGCTGGAGGATAGAGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC 3000
DB 2941 TTCTGTCTCTGCTGGAGGATAGAGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC 3000
QY 3001 CAGGTGACTGACTTGCAGTGTGTGCTGTGTGCAGAAAGATGTGTGGCAGTCTGAACATC 3060
DB 3001 CAGGTGACTGACTTGCAGTGTGTGCTGTGTGCAGAAAGATGTGTGGCAGTCTGAACATC 3060
QY 3061 TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC 3120
DB 3061 TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC 3120
QY 3121 GCTAGGCTGCCGGGACGCTGTACCTGAGACAGAGCTGTATGTTAGCTGCACCTGTG 3180
DB 3121 GCTAGGCTGCCGGGACGCTGTACCTGAGACAGAGCTGTATGTTAGCTGCACCTGTG 3180
QY 3181 GAGCAACATGGGCGTGTCTGCAGAACTGCGTGGTCTTGGCTGTACTGCTGTGTGTGC 3240
DB 3181 GAGCAACATGGGCGTGTCTGCAGAACTGCGTGGTCTTGGCTGTACTGCTGTGTGTGC 3240
QY 3241 GCGTGTGTTCTGGGGTGAGTTCGTGAAATGATGTGTGTGTCAGGGCCATCAGCAAGGGTAA 3300
DB 3241 GCGTGTGTTCTGGGGTGAGTTCGTGAAATGATGTGTGTGTCAGGGCCATCAGCAAGGGTAA 3300
QY 3301 GAACGAGCGGGCGGCTGAGTTCAGCCCTGTAATCCAGCCCTTTGGGAGGCGGAGCA 3360
DB 3301 GAACGAGCGGGCGGCTGAGTTCAGCCCTGTAATCCAGCCCTTTGGGAGGCGGAGCA 3360
QY 3361 GCGGATCACTGAGGTGCGGAGATCGAGGCCAGCTGACCAACATGGAGAACCCCGTCT 3420
DB 3361 GCGGATCACTGAGGTGCGGAGATCGAGGCCAGCTGACCAACATGGAGAACCCCGTCT 3420
QY 3421 CTACTAAAAATACAAAAATTAGCTGTGTGTGGCGCGTGCCTGTAATCCAGCTACTC 3480
DB 3421 CTACTAAAAATACAAAAATTAGCTGTGTGTGGCGCGTGCCTGTAATCCAGCTACTC 3480
QY 3481 GGGAGCTGGGCGAGAAAAATCCGTTGAACCCGGAGGTGGAGTTGCGGTGAGCCGAGA 3540
DB 3481 GGGAGCTGGGCGAGAAAAATCCGTTGAACCCGGAGGTGGAGTTGCGGTGAGCCGAGA 3540
QY 3541 TCGCGCATTTGCACTTCAGCCTGGGCAACAAGACGAAACTCCGCTCTCGAAAGAAAAAA 3600
DB 3541 TCGCGCATTTGCACTTCAGCCTGGGCAACAAGACGAAACTCCGCTCTCGAAAGAAAAAA 3600
QY 3601 GAAAAAAGGTAAAGAACCAAGTGAATGGGCACGGAGGACTGATGATGGAGTGGGC 3660
DB 3601 GAAAAAAGGTAAAGAACCAAGTGAATGGGCACGGAGGACTGATGATGGAGTGGGC 3660
QY 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGATTTGACAGGATTTGAGAGGCAT 3720
DB 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGATTTGACAGGATTTGAGAGGCAT 3720
QY 3721 GTTTTCACTGAGAAATTCAGAAACCTTAGGCTGTCTTTCCCTCCATGTGGCCCTTAAG 3780
DB 3721 GTTTTCACTGAGAAATTCAGAAACCTTAGGCTGTCTTTCCCTCCATGTGGCCCTTAAG 3780
QY 3781 CTGAGCCCTTCTTTCTCGGTCTGCTTTTCGGAACCCCTAGCTCCGCCCATGAGCTCTGACC 3840

DB 3781 CTGAGCCCTTCTTTCTCGGTCTGCTTTGGAACCCCTAGCTCCGCCCATGAGCTCTGACC 3840
QY 3841 CCACCTCCTTTCTCTCAACACAGCCCTAGGCCAGACTCTTAGTGAGACCCCGCCTTAAGSCCA 3900
DB 3841 CCACCTCCTTTCTCTCAACACAGCCCTAGGCCAGACTCTTAGTGAGACCCCGCCTTAAGSCCA 3900
QY 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTTCTGTGGGTACCTTTAGAAACCCCTTTCAA 3960
DB 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTTCTGTGGGTACCTTTAGAAACCCCTTTCAA 3960
QY 3961 AGTCAGAGCTTTTCTTTTCTTTTGGAGACAGCTTTGCTCTCTCTCCAGGCTGGAG 4020
DB 3961 AGTCAGAGCTTTTCTTTTCTTTTGGAGACAGCTTTGCTCTCTCTCTCCAGGCTGGAG 4020
QY 4021 TGCAGTGGCGTGATCTCGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGTGATTTCTGT 4080
DB 4021 TGCAGTGGCGTGATCTCGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGTGATTTCTGT 4080
QY 4081 GCCTCCACCTCTCTGAGTAGTGGGATTTACAGGTGCGCGGCCACCAACGCTGGCTAAATTTT 4140
DB 4081 GCCTCCACCTCTCTGAGTAGTGGGATTTACAGGTGCGCGGCCACCAACGCTGGCTAAATTTT 4140
QY 4141 GTGCTTTTATGATAGACAGGGTTTCACTTTGTTGGCCAGGCTGGTCTCAAACTCCCAACC 4200
DB 4141 GTGCTTTTATGATAGACAGGGTTTCACTTTGTTGGCCAGGCTGGTCTCAAACTCCCAACC 4200
QY 4201 TCAGGTGATCCGCCACCTCGGCTCCAGAGTCTGGGGTTACAGGCGTGAGCCACCGC 4260
DB 4201 TCAGGTGATCCGCCACCTCGGCTCCAGAGTCTGGGGTTACAGGCGTGAGCCACCGC 4260
QY 4261 CCCAGCCCAAGTCAAGAGCTCTTTATAGAGACTCTAAACATGTAACCTGACCTGGGCC 4320
DB 4261 CCCAGCCCAAGTCAAGAGCTCTTTATAGAGACTCTAAACATGTAACCTGACCTGGGCC 4320
QY 4321 CTAACTAAGTCAATTTCCAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380
DB 4321 CTAACTAAGTCAATTTCCAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380
QY 4381 TGACCCCACTTCTTGAGACAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440
DB 4381 TGACCCCACTTCTTGAGACAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440
QY 4441 GCTCCAGCCCCCAGCTTTGGCACTACCCCTGAGCTTTGTCCAGGAATCTGTACCCAAT 4500
DB 4441 GCTCCAGCCCCCAGCTTTGGCACTACCCCTGAGCTTTGTCCAGGAATCTGTACCCAAT 4500
QY 4501 TTTACCCCTCACATGTATGTCTAGCCAAATTCAGGAATCTGTGAGGTCCAGTTAGAGTCCA 4560
DB 4501 TTTACCCCTCACATGTATGTCTAGCCAAATTCAGGAATCTGTGAGGTCCAGTTAGAGTCCA 4560
QY 4561 GTAAACCTTACCTGAGCTGGCTCTGCTCTTGAGCTTTGAGCCCTGGGCTTGAGAGGTGCCA 4620
DB 4561 GTAAACCTTACCTGAGCTGGCTCTGCTCTTGAGCTTTGAGCCCTGGGCTTGAGAGGTGCCA 4620
QY 4621 CTCTTATTTCTCCAGGCGCTGCCCTGCCCTCAGCATGTTCAGACACCCACCTCTAGCT 4680
DB 4621 CTCTTATTTCTCCAGGCGCTGCCCTGCCCTCAGCATGTTCAGACACCCACCTCTAGCT 4680
QY 4681 GGTCTGGCTCTTTGAGTGTGAACACCCCAAGCCCAAGCCCGCTCTGAGCCCGGCC 4740
DB 4681 GGTCTGGCTCTTTGAGTGTGAACACCCCAAGCCCAAGCCCGCTCTGAGCCCGGCC 4740
QY 4741 CAAACCAATTTCCGTTCCAGAGACTGTTCTCGCCCAATGATGTTTCTGTGACCCACC 4800
DB 4741 CAAACCAATTTCCGTTCCAGAGACTGTTCTCGCCCAATGATGTTTCTGTGACCCACC 4800
QY 4801 CCTCTAACACCGTGCCTCTGGGAGCAACAGGAGCTGGGAGCTGGGCGCGGGAAGACG 4860
DB 4801 CCTCTAACACCGTGCCTCTGGGAGCAACAGGAGCTGGGAGCTGGGCGCGGGAAGACG 4860
QY 4861 CCGGCTGGAGTGAACAGAGAGCCGATCATGATGATCCGATCGGATATGACACCC 4920

Db 4861 CCCGTCGGATGACAGCAGCAGCCGATCATCAATGGATCCGATCGATATGACACACC 4920
Qy 4921 AGCGTGGCAGCGCGCTGTGTCTAAGGCCCAACAGCTTACTGCGGGCGGTGTGG 4980
Db 4921 AGCGTGGCAGCGCGCTGTGTCTAAGGCCCAACAGCTTACTGCGGGCGGTGTGG 4980
Qy 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 3
ADR72623
ID ADR72623 standard; DNA; 11570 BP.
AC ADR72623;
XX 02-DEC-2004 (first entry)
XX Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.
XX kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;
KW cancer metastasis; chemotherapy; human; serine protease;
KW chromosome 19q13.4; KLK5; ds; gene.
XX Homo sapiens.
XX

Key Location/Qualifiers
FT 2221..11247
FT CDS
FT /*tag= b
FT /product= "Human renal cell carcinoma-related kallikrein
FT 5 (KLK5) protein"
FT 2221..2293
FT /*tag= a
FT /number= 1
FT 2294..4761
FT /*tag= c
FT /number= 1
FT 4762..5023
FT /*tag= d
FT /number= 2
FT 5024..5762
FT /*tag= e
FT /number= 2
FT 5763..6019
FT /*tag= f
FT /number= 3
FT 6020..6104
FT /*tag= g
FT /number= 3
FT 6105..6238
FT /*tag= h
FT /number= 4
FT 6239..11091
FT /*tag= i
FT /number= 4
FT 11092..11247
FT /*tag= j
FT /number= 5

WO2004077060-A2.
XX
PN 10-SEP-2004.
XX
XX 26-FEB-2004; 2004WO-CA000280.
XX
XX 27-FEB-2003; 2003US-0451382P.
XX
XX (MOUN) MOUNT SINAI HOSPITAL.
XX
XX Diamandis EP, Petraki CD;
XX
XX WPI; 2004-662077/64.

DR P-PSDB; ADR72621.
XX
FT Detecting kallikrein polypeptides/polynucleotides associated with renal
PT cell carcinoma in a patient, for diagnosing/treating the disease,
PT comprises detecting /identifying kallikrein polypeptides/polynucleotides
XX in a sample.
PS
XX Example 1; SEQ ID NO 3; 53pp; English.
CC
CC The invention relates to a novel method for detecting kallikrein
CC polypeptides, or the polynucleotides encoding them, associated with renal
CC cell carcinoma. The method comprises obtaining a sample from a patient
CC and detecting kallikrein polypeptides, or their encoding polynucleotides,
CC where the kallikrein polypeptides are selected from the group consisting
CC of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The
CC detected amounts of the kallikrein polypeptides are compared to standard
CC amounts. The molecules of the invention demonstrate cytostatic activity
CC whilst the methods and kit may be useful for detecting, characterising,
CC preventing and treating renal cell carcinoma. Furthermore, the methods
CC may be useful for evaluating the probability of the presence of malignant
CC or pre-malignant cells and for detecting and quantitating tumour growth
CC and cancer metastasis. Finally, the methods may be utilised to confirm
CC the absence or removal of all tumour tissue following surgery, cancer
CC chemotherapy and/or radiation therapy and to monitor cancer chemotherapy
CC and tumour re-appearance. The current sequence is that of the human
CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the
CC invention which encodes a secreted serine protease and is located at
CC chromosome 19q13.4.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 5000; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGCCACAGTGAAGGCAAGAGAGAGTGTGAGAGTCCCTCTCAAGTGGCTTGAGTC 60
Db 1 GGGCCACAGTGAAGGCAAGAGAGAGTGTGAGAGTCCCTCTCTGCAAGTGGCTTGAGTC 60
Qy 61 TCCCTGCTCTAAATGTCAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTGCTCTAAATGTCAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 AAGAAAG 180
Db 121 AAGAAAG 180
Qy 181 ACAGAGAGCCTGGGACACAGCGGACACACAGAGTCCAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCCTGGGACACAGCGGACACACAGAGTCCAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 ACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGAGATTAAACAGAGTCCACAGATACACGC 300
Db 241 ACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGAGATTAAACAGAGTCCACAGATACACGC 300
Qy 301 AAAGGGGACAGACAGAGTTTTCAGGGTGTGTCTATGATCATCTTCTTTTTTTTTTTT 360
Db 301 AAAGGGGACAGACAGAGTTTTCAGGGTGTGTCTATGATCATCTTCTTTTTTTTTTTT 360
Qy 361 TTTTCTTTTTTTTTTTTGGAGCGGAGTCTCGCTCTGTGCGCCCGAGGAGTGCAGTGGCGG 420
Db 361 TTTTCTTTTTTTTTTTTGGAGCGGAGTCTCGCTCTGTGCGCCCGAGGAGTGCAGTGGCGG 420
Qy 421 GATCTCGGCTCACTGCAAGCTCCCGGCTTCCAGGCAATTCCTCTGCTCAGGCTC 480
Db 421 GATCTCGGCTCACTGCAAGCTCCCGGCTTCCAGGCAATTCCTCTGCTCAGGCTC 480
Qy 481 CCAAGTAGCTGGGACTACAGCGCCCGCCACTACAGCGCCCGGCTAATTTTTTGTATTTTA 540
Db 481 CCAAGTAGCTGGGACTACAGCGCCCGCCACTACAGCGCCCGGCTAATTTTTTGTATTTTA 540
Qy 541 GTAGAGACGGGGTTTACCGCTTTTAGCGGGAGTGGCTCGATCTCTCTGACCTCGTGATCC 600
Db 541 GTAGAGACGGGGTTTACCGCTTTTAGCGGGAGTGGCTCGATCTCTCTGACCTCGTGATCC 600

Db 541 GTAGAGAGGGGTTTACCGTTTTAGCCGGGATGGCCCTCGATCTCTGACCTCGTGATCC 600
Qy
G 601 GCCCGCTCGGCCTCCCAAGTGTCTGGATTATACAGGCGTGAGCCACCGCCCGCCCATG 660
Db
G 601 GCCCGCTCGGCCTCCCAAGTGTCTGGATTATACAGGCGTGAGCCACCGCCCGCCCATG 660
Qy
A 661 ATCATCTTCTGACTATGCTGATGTGACAAGTACCTTAAAGCCCATCAGACTCTACCCCTTTA 720
Db
A 661 ATCATCTTCTGACTATGCTGATGTGACAAGTACCTTAAAGCCCATCAGACTCTACCCCTTTA 720
Qy
A 721 AATATGACAGTTTGGCCAGGACCGTGCTCATGCCCTGTAAATTCACGACTTTGGGAGGC 780
Db
A 721 AATATGACAGTTTGGCCAGGACCGTGCTCATGCCCTGTAAATTCACGACTTTGGGAGGC 780
Qy
A 781 AGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACACGAGCTTGGCCCAACATGTTGAAA 840
Db
A 781 AGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACACGAGCTTGGCCCAACATGTTGAAA 840
Qy
C 841 CTCGTCTTTTACTAAAAAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db
C 841 CTCGTCTTTTACTAAAAAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Qy
A 901 ACCTGTAAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAACCCCTGGAGGC 960
Db
A 901 ACCTGTAAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAACCCCTGGAGGC 960
Qy
G 961 GAGGTTGAGTGGCCGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGAGCAAGACT 1020
Db
G 961 GAGGTTGAGTGGCCGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGAGCAAGACT 1020
Qy
C 1021 CTGTCTCAAAATAAATAAACAACAACAAGCAGTTTGTGTACCTTGTATATATCT 1080
Db
C 1021 CTGTCTCAAAATAAATAAACAACAACAAGCAGTTTGTGTACCTTGTATATATCT 1080
Qy
A 1081 AAAAAAATAATGCTGTCAACAATAATAGAGCAAGTGAATTAAGGAAATAAATGGGCCA 1140
Db
A 1081 AAAAAAATAATGCTGTCAACAATAATAGAGCAAGTGAATTAAGGAAATAAATGGGCCA 1140
Qy
A 1141 AGAAGCTTAAGGTATATTTGCAAAATCATTTGCAAACTTTTAAAGGAGGAGGAGGAGGAG 1200
Db
A 1141 AGAAGCTTAAGGTATATTTGCAAAATCATTTGCAAACTTTTAAAGGAGGAGGAGGAGGAG 1200
Qy
G 1201 GGATAGAAAGACAGGGAGGAAACAGGAGACAGAAAACCTGTGGCCCAAGGAGAAACAAA 1260
Db
G 1201 GGATAGAAAGACAGGGAGGAAACAGGAGACAGAAAACCTGTGGCCCAAGGAGAAACAAA 1260
Qy
A 1261 ACAAGGCTCTTAAGACAGACAGGAGGAGAGAGAGAGAGTGTGAGACAGACAGAG 1320
Db
A 1261 ACAAGGCTCTTAAGACAGACAGGAGGAGAGAGAGAGAGTGTGAGACAGACAGAG 1320
Qy
A 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAGGAGGAGGAGGAGGAG 1380
Db
A 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAGGAGGAGGAGGAGGAG 1380
Qy
G 1381 GAGAGAGGGGTGGAGAGAGACACAGATATTGAGAGAGACTCAGAAAAGATAGCCGAGGGA 1440
Db
G 1381 GAGAGAGGGGTGGAGAGAGACACAGATATTGAGAGAGACTCAGAAAAGATAGCCGAGGGA 1440
Qy
G 1441 GAACCAAGAGAGATGAAGAGAGACTCTGAGAAAACCAAGAGACAAAGATGGAAGAGG 1500
Db
G 1441 GAACCAAGAGAGATGAAGAGAGACTCTGAGAAAACCAAGAGACAAAGATGGAAGAGG 1500
Qy
A 1501 AGTATCCAGGGTGAACAGAGAGTGGTGAATGACAAAATGCCAGAGAGGAGGAGGAGGAG 1560
Db
A 1501 AGTATCCAGGGTGAACAGAGAGTGGTGAATGACAAAATGCCAGAGAGGAGGAGGAGGAG 1560
Qy
T 1561 TCCAGGCGCCAAAGTATGATGACAGAGTTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db
T 1561 TCCAGGCGCCAAAGTATGATGACAGAGTTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Qy
A 1621 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Db
A 1621 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680

Qy 1681 GTAGGAGGTGACATTCGGACTGGGTGGGGGGTGTCTTGGGGGTGGAGATAGGGGGAGCA 1740
Db
G 1681 GTAGGAGGTGACATTCGGACTGGGTGGGGGGTGTCTTGGGGGTGGAGATAGGGGGAGCA 1740
Qy
G 1741 GGAGGAGCTATTGCTAAGGCCCGGATAGGCACTCTCATTTGCCCGGGAAATGTCCCCAGGGAG 1800
Db
G 1741 GGAGGAGCTATTGCTAAGGCCCGGATAGGCACTCTCATTTGCCCGGGAAATGTCCCCAGGGAG 1800
Qy
C 1801 CAGTGGGTGGTTATACTCAGGCCCGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db
C 1801 CAGTGGGTGGTTATACTCAGGCCCGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Qy
A 1861 AGGCACAGGCTCTGAGAAAGTCTGGGCTGAGCTGGGAGCAAAATCCCCACCCCTTACTCTGG 1920
Db
A 1861 AGGCACAGGCTCTGAGAAAGTCTGGGCTGAGCTGGGAGCAAAATCCCCACCCCTTACTCTGG 1920
Qy
G 1921 GGACAGGGCAAGTGTAGAGACTGTGTAGGGTGTCTCAGAGGAGGAGGAGGAGGAGGAGGAG 1980
Db
G 1921 GGACAGGGCAAGTGTAGAGACTGTGTAGGGTGTCTCAGAGGAGGAGGAGGAGGAGGAGGAG 1980
Qy
G 1981 GTGGGTCTGCACCCACATCTTTCTCTGTCCCTCTCTGCTCTGCTGTCTGAGAGGCTGTAGA 2040
Db
G 1981 GTGGGTCTGCACCCACATCTTTCTCTGTCCCTCTCTGCTCTGCTGTCTGAGAGGCTGTAGA 2040
Qy
C 2041 CTCTATCTTCTGAATTTCTATAGTGTCTGGGTCTCAGCGCAGTGGCCGATGGTGGCCCGTC 2100
Db
C 2041 CTCTATCTTCTGAATTTCTATAGTGTCTGGGTCTCAGCGCAGTGGCCGATGGTGGCCCGTC 2100
Qy
C 2101 CTTGTGGTCTCTCTCACTGGGGAATAAGGTAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db
C 2101 CTTGTGGTCTCTCTCACTGGGGAATAAGGTAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Qy
G 2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTTCCCATTCAGAGTGCAGCGGCC 2220
Db
G 2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTTCCCATTCAGAGTGCAGCGGCC 2220
Qy
A 2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGTGCTCTGTATCAGAGCTTGCTTT 2280
Db
A 2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGTGCTCTGTATCAGAGCTTGCTTT 2280
Qy
C 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGGAGGGTGTGGGATTTGGGAGGAGTGT 2340
Db
C 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGGAGGGTGTGGGATTTGGGAGGAGTGT 2340
Qy
C 2341 CTCTGGGCACTAGAGCGCTTCCCTGGGGAACCTGTGTGAGCCTGGGAGTGCATCCGG 2400
Db
C 2341 CTCTGGGCACTAGAGCGCTTCCCTGGGGAACCTGTGTGAGCCTGGGAGTGCATCCGG 2400
Qy
G 2401 GACCGGTGAATGTGAGTCTCTGTCTGTACTTGTGGTGTGGCATCGTATGTGGCCCTGT 2460
Db
G 2401 GACCGGTGAATGTGAGTCTCTGTCTGTACTTGTGGTGTGGCATCGTATGTGGCCCTGT 2460
Qy
G 2461 GACTGCCACGGTGTGTCTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCGG 2520
Db
G 2461 GACTGCCACGGTGTGTCTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCGG 2520
Qy
C 2521 CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db
C 2521 CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Qy
A 2581 ATTGTGTGGGTCCACAGCTGTGGGTGAATGCAATGACACTGGGGGTGTTCACGTGT 2640
Db
A 2581 ATTGTGTGGGTCCACAGCTGTGGGTGAATGCAATGACACTGGGGGTGTTCACGTGT 2640
Qy
G 2641 GTGTTTGGCTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGT 2700
Db
G 2641 GTGTTTGGCTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGGTGT 2700
Qy
T 2701 TCCCTGAGGTCGGGATTCGCTGCAACAAAGTGTTCATCACCATGGAAGCTGTGACT 2760
Db
T 2701 TCCCTGAGGTCGGGATTCGCTGCAACAAAGTGTTCATCACCATGGAAGCTGTGACT 2760

QY 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTGTGGCTGAGTGTGACCTATGATGCCGTA 2820
Db 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTGTGGCTGAGTGTGACCTATGATGCCGTA 2820
QY 2821 TTTGTGACCGGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG 2880
Db 2821 TTTGTGACCGGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG 2880
QY 2881 TCTGTGTGAGCGCGTGTAAATGCTACTGTATGTGTGATGTGGACCTGTGTCTGGAGT 2940
Db 2881 TCTGTGTGAGCGCGTGTAAATGCTACTGTATGTGTGATGTGGACCTGTGTCTGGAGT 2940
QY 2941 TTTCTGTCTCTCCCTGGAGGATAGAGGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC 3000
Db 2941 TTTCTGTCTCTCCCTGGAGGATAGAGGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC 3000
QY 3001 CAGGTGACTGACTGTGCACTGTGCTGTGTGCGAGAGAGTATGTGGCAGTCTGAACATC 3060
Db 3001 CAGGTGACTGACTGTGCACTGTGCTGTGTGCGAGAGAGTATGTGGCAGTCTGAACATC 3060
QY 3061 TGTGCACACAGCGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGGTGTGGATCCC 3120
Db 3061 TGTGCACACAGCGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGGTGTGGATCCC 3120
QY 3121 GCTAGGCTGCCGCGAGCGTGTACTTACCTGGAGACAGAGCTGTATGTGTAGCTGCACTGTG 3180
Db 3121 GCTAGGCTGCCGCGAGCGTGTACTTACCTGGAGACAGAGCTGTATGTGTAGCTGCACTGTG 3180
QY 3181 GAGCAACATAGGGCGTGTCTGAGAACTGCGTGGCTGTGGCTGTACTGCTGTGTGTC 3240
Db 3181 GAGCAACATAGGGCGTGTCTGAGAACTGCGTGGCTGTGGCTGTACTGCTGTGTGTC 3240
QY 3241 GCGTGGTCTTGGGGTGTGCTGTAATGATGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3300
Db 3241 GCGTGGTCTTGGGGTGTGCTGTAATGATGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3300
QY 3301 GAACAGCGCGCGCGTGTGCTCAAGCTGTAAATCCAGCGCTTGGGGCGCGAGGCA 3360
Db 3301 GAACAGCGCGCGCGTGTGCTCAAGCTGTAAATCCAGCGCTTGGGGCGCGAGGCA 3360
QY 3361 GCGGGATCACCTGAGGTGCGGAGATCGAGGCCAGCTGACCAACATGAGAGAACCCGCTCT 3420
Db 3361 GCGGGATCACCTGAGGTGCGGAGATCGAGGCCAGCTGACCAACATGAGAGAACCCGCTCT 3420
QY 3421 CTACTAAATAATACAAAAATTTAGCTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3480
Db 3421 CTACTAAATAATACAAAAATTTAGCTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3480
QY 3481 GGGAGACTGGGGCAGAAAAATCGCTTGAACCCGGAGGTGGAGCTTGGGTGAGCCGAGA 3540
Db 3481 GGGAGACTGGGGCAGAAAAATCGCTTGAACCCGGAGGTGGAGCTTGGGTGAGCCGAGA 3540
QY 3541 TCGCGCCATTGCACTCCAGCTCGGGCAACAGAGCGAAATCTCCGCTCGAAAGAAAAAA 3600
Db 3541 TCGCGCCATTGCACTCCAGCTCGGGCAACAGAGCGAAATCTCCGCTCGAAAGAAAAAA 3600
QY 3601 GAAAAAAGGTGTAAACAGTGAATGGGCAACGGAGGACTGTATGATGGAGTGGGGC 3660
Db 3601 GAAAAAAGGTGTAAACAGTGAATGGGCAACGGAGGACTGTATGATGGAGTGGGGC 3660
QY 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGGATGACAGATTGAGAGGCGAT 3720
Db 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGGATGACAGATTGAGAGGCGAT 3720
QY 3721 GTTTTTCATCTGAGAAATCAGAAATCCTAGGCTGTCTTTCCCTCCATGTGGCCCTTAAG 3780
Db 3721 GTTTTTCATCTGAGAAATCAGAAATCCTAGGCTGTCTTTCCCTCCATGTGGCCCTTAAG 3780
QY 3781 CTGAGCCCTCTTTCTGGTCTCTGTTTGGAAACCTTAGCTCCGCCCATGAGCTCTGACC 3840
Db 3781 CTGAGCCCTCTTTCTGGTCTCTGTTTGGAAACCTTAGCTCCGCCCATGAGCTCTGACC 3840
QY 3841 CCACCTCTCTTCTCAACACGCGCCCTTAGGCCAGACTCTAGTGGACCCCGCTTAAGGCCA 3900

Db 3841 CCACCTCTCTTCTCAACACGCGCCCTTAGGCCAGACTCTAGTGGACCCCGCTTAAGGCCA 3900
QY 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTCTGTGGTACCTTTCTAGAACCCCTTTCAA 3960
Db 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTCTGTGGTACCTTTCTAGAACCCCTTTCAA 3960
QY 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGCTGGAG 4020
Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGCTGGAG 4020
QY 4021 TGCAGTGGCGTGTACTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
Db 4021 TGCAGTGGCGTGTACTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
QY 4081 GCCTCCACTCTCTGAGTAGCTGGGATTACAGGTGCGCGCACACGCTTGCTTAATTTTT 4140
Db 4081 GCCTCCACTCTCTGAGTAGCTGGGATTACAGGTGCGCGCACACGCTTGCTTAATTTTT 4140
QY 4141 GTGTCTTTAGTAGAGACAGGGTTTTCACCTTTGTTGGCCAGGCTGTCTCAAACTCCCAACC 4200
Db 4141 GTGTCTTTAGTAGAGACAGGGTTTTCACCTTTGTTGGCCAGGCTGTCTCAAACTCCCAACC 4200
QY 4201 TCAGGTGATCCGCCACTCTCGGCTCCAGAGTGTCTGGGTTACAGCGTGGAGCACCGC 4260
Db 4201 TCAGGTGATCCGCCACTCTCGGCTCCAGAGTGTCTGGGTTACAGCGTGGAGCACCGC 4260
QY 4261 CCCAGGCCAAAGTCAAGCTCTTTATPAGGAGACTCTAAACATGTAACTGACCTTGSCC 4320
Db 4261 CCCAGGCCAAAGTCAAGCTCTTTATPAGGAGACTCTAAACATGTAACTGACCTTGSCC 4320
QY 4321 CTAACTAAGTCAATTCAAAACCCCTTCTGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 4380
Db 4321 CTAACTAAGTCAATTCAAAACCCCTTCTGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAG 4380
QY 4381 TGACCCCACTCTTTGAGACAGTTTCCATCCCTAAAGCCCTGCTCTCCCTCCATCCCGAG 4440
Db 4381 TGACCCCACTCTTTGAGACAGTTTCCATCCCTAAAGCCCTGCTCTCCCTCCATCCCGAG 4440
QY 4441 GCTCCAGCCCCCAGCTTTGGCACTACCCCTGAGCTTTGTCCAGGAATCTCTGTACCCAAT 4500
Db 4441 GCTCCAGCCCCCAGCTTTGGCACTACCCCTGAGCTTTGTCCAGGAATCTCTGTACCCAAT 4500
QY 4501 TTTACCTTCACATGTAGTCTAGCCAAATTCAGGAATCTGTGAGGTCCAGTTAGAGTCCA 4560
Db 4501 TTTACCTTCACATGTAGTCTAGCCAAATTCAGGAATCTGTGAGGTCCAGTTAGAGTCCA 4560
QY 4561 GTAACTTACCTGAGCTGGGCTCTGTCTTTGAGCTTTGAGCTGGGCTTGGAGAGTGGCA 4620
Db 4561 GTAACTTACCTGAGCTGGGCTCTGTCTTTGAGCTTTGAGCTGGGCTTGGAGAGTGGCA 4620
QY 4621 CTCTTATTTCTCCAGGCGCTTCCCTTCCCTCAGCATGTCCAGACACCCCTCTAGCT 4680
Db 4621 CTCTTATTTCTCCAGGCGCTTCCCTTCCCTCAGCATGTCCAGACACCCCTCTAGCT 4680
QY 4681 GGTCTGGCTCTTTGAGTCTGAAACCCACCCCGCAAGCCCGCTCTGAGCCCGGCC 4740
Db 4681 GGTCTGGCTCTTTGAGTCTGAAACCCACCCCGCAAGCCCGCTCTGAGCCCGGCC 4740
QY 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTGTGACCAAC 4800
Db 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTGTGACCAAC 4800
QY 4801 CCTCTAACACGCTGCTTGGGCAACACAGACCTTGGGAGCTGGGGCGGGGGAAGCG 4860
Db 4801 CCTCTAACACGCTGCTTGGGCAACACAGACCTTGGGAGCTGGGGCGGGGGAAGCG 4860
QY 4861 CCCGGTTCGATGACAGCAGCGCGCATCATCAATGGATCCGACTTCGATATGACACCC 4920
Db 4861 CCCGGTTCGATGACAGCAGCGCGCATCATCAATGGATCCGACTTCGATATGACACCC 4920
QY 4921 AGCGGTGGAGGCGCGCTTGTGTTAAGGCCCAACAGCTCTACTTGGGGGCGGTGTGG 4980

Db 4921 AGCGTGGAGGGCGCGCTGTTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980
Qy 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 4
ADR72875
ID ADR72875 standard; DNA; 11570 BP.
XX ADR72875;
XX 02-DEC-2004 (first entry)
XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.
XX kallikrein 5; tumour marker; ovarian cancer;
XX epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;
KW KLK5; ds; gene.
XX Homo sapiens.

XX Location/Qualifiers
FH Key 2221..11247
FT CDS /tag= b
FT /product= "Human ovarian cancer-related tumour marker
FT kallikrein 5 (hk5) protein"
FT exon 2221..2293
FT /tag= a
FT /number= 1
FT intron 2294..4761
FT /tag= c
FT /number= 1
FT exon 4762..5023
FT /tag= d
FT /number= 2
FT intron 5024..5762
FT /tag= e
FT /number= 2
FT exon 5763..6019
FT /tag= f
FT /number= 3
FT intron 6020..6104
FT /tag= g
FT /number= 3
FT exon 6105..6238
FT /tag= h
FT /number= 4
FT exon 6239..11091
FT /tag= i
FT /number= 4
FT exon 11092..11247
FT /tag= j
FT /number= 5

W02004075713-A2.
10-SEP-2004.
26-FEB-2004; 2004WO-CA000281.
26-FEB-2003; 2003US-0450406P.
(MOUN) MOUNT SINAI HOSPITAL.
Diamandis EP;
WPI: 2004-661815/64.
P-PSDB; ADR72873.
Kallikrein markers detection method for detecting ovarian cancer in
patient, involves detecting kallikrein markers and CA125 in sample

PT obtained from patient, and comparing detected amounts with standard
XX amounts.
PS Example 2; SEQ ID NO 5; 102pp; English.
XX
CC The invention relates to a novel method for detecting a plurality of
CC kallikrein markers associated with ovarian cancer. The method comprises
CC obtaining a sample from a patient and detecting in the sample a plurality
CC of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
CC the kallikrein markers are selected from the group consisting of
CC kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
CC kallikrein 11. The detected amounts of the kallikrein markers are
CC compared with standard amounts. The method of the invention may be useful
CC for detecting kallikrein markers associated with ovarian cancer in a
CC patient and thus for detecting ovarian cancer, particularly epithelial
CC ovarian carcinoma. The current sequence is that of the human ovarian
CC cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention
CC which encodes a secreted serine protease and is located at chromosome
CC 19q13.4.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 5000; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGCCACAGAGTGAAGGCAAGAGAGAGTGTAGAGCTCCCTCTGCAAAAGTGGCTTGAATC 60
Db 1 GGGCCACAGAGTGAAGGCAAGAGAGAGTGTAGAGCTCCCTCTGCAAAAGTGGCTTGAATC 60
Qy 61 TCCCTTGCCTAAATGCAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTTGCCTAAATGCAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAACTACAGAAACACAGAGAGAGACAC 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAACTACAGAAACACAGAGAGAGACAC 180
Qy 181 ACAGAGAGCTGGGACACAGAGGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGAGGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 ACACAAATGGAGACACAGAGGTTGAAGAAAGAGAGATTAACAGAGTCCAGATACACGC 300
Db 241 ACACAAATGGAGACACAGAGGTTGAAGAAAGAGAGATTAACAGAGTCCAGATACACGC 300
Qy 301 AAAGGGGCAAGACACAGTTTTTCAGGGTGTGTATGATCATCTCTTTTTTTTTTTTTT 360
Db 301 AAAGGGGCAAGACACAGTTTTTCAGGGTGTGTATGATCATCTCTTTTTTTTTTTTTT 360
Qy 361 TTTTTTTTTTTTTTTTTCAGAGCGGAGTCTCGCTCTGCGCCAGCTGGAGTGGAGTGGCGG 420
Db 361 TTTTTTTTTTTTTTTTTCAGAGCGGAGTCTCGCTCTGCGCCAGCTGGAGTGGAGTGGCGG 420
Qy 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 480
Db 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 480
Qy 481 CCAAGTAGCTGGAGCTACAGGGCGCGCCCACTACAGGGCGCGCTAAATTTTTTGTATTTTA 540
Db 481 CCAAGTAGCTGGAGCTACAGGGCGCGCGCCCACTACAGGGCGCGCTAAATTTTTTGTATTTTA 540
Qy 541 GTAGAGACGGGGTTTACCGTTTTTACCGGGTGGCTCGATCTCCGAGCTCGTATCC 600
Db 541 GTAGAGACGGGGTTTACCGTTTTTACCGGGTGGCTCGATCTCCGAGCTCGTATCC 600
Qy 601 GCCCGCTCGGCTCCCAAGTCTGGGATTACAGGCGTGGAGCCACCGCGCCCGCCCATG 660
Db 601 GCCCGCTCGGCTCCCAAGTCTGGGATTACAGGCGTGGAGCCACCGCGCCCGCCCATG 660
Qy 661 ATCATCTTCTGATCTATGCTGATGACAAAGTACCTAAAGCCATCAGACTCTACCCCTTA 720
Db 661 ATCATCTTCTGATCTATGCTGATGACAAAGTACCTAAAGCCATCAGACTCTACCCCTTA 720

QY 721 AATATGAGTTTGGCCAGGACCGTGGCTCATGCTGTAAATCCACGACTTTGGGAGGC 780
DB |||||||
QY 721 AATATGAGTTTGGCCAGGACCGTGGCTCATGCTGTAAATCCACGACTTTGGGAGGC 780
DB |||||||
QY 781 AGAGGTGGGTGAATCACTTTGAGGCGCAGAGTTTGGAGACCGCTGGCCAAACATGGTGA 840
DB |||||||
QY 781 AGAGGTGGGTGAATCACTTTGAGGCGCAGAGTTTGGAGACCGCTGGCCAAACATGGTGA 840
DB |||||||
QY 841 CTCTGTCTTTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCTGGGGCAC 900
DB |||||||
QY 841 CTCTGTCTTTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCTGGGGCAC 900
DB |||||||
QY 901 ACCTGTAATCCAGCTATCTGAGGCTGAGGACGAGAGTCACTTCAACCTTGAGGGCG 960
DB |||||||
QY 901 ACCTGTAATCCAGCTATCTGAGGCTGAGGACGAGAGTCACTTCAACCTTGAGGGCG 960
DB |||||||
QY 961 GAGTTGAGTGGGCGGAGATCAATCACCGCCCTCCAGCCCTGGCGCACGAGCAAGACT 1020
DB |||||||
QY 961 GAGTTGAGTGGGCGGAGATCAATCACCGCCCTCCAGCCCTGGCGCACGAGCAAGACT 1020
DB |||||||
QY 1021 CTGTCTCAATTAATAATAAACAACGAAACAGCAGTCTTTGTGTACTTGTATATCT 1080
DB |||||||
QY 1021 CTGTCTCAATTAATAATAAACAACGAAACAGCAGTCTTTGTGTACTTGTATATCT 1080
DB |||||||
QY 1081 AAAAAAAAAATGCTCAACAAATAGACGAGTGAATTAAGGAAATAAATGGGCCA 1140
DB |||||||
QY 1081 AAAAAAAAAATGCTCAACAAATAGACGAGTGAATTAAGGAAATAAATGGGCCA 1140
DB |||||||
QY 1141 AGAACTCTAAGGTATATTTGACAAATCAATTCAGAACTTTTAAAAAGAAAGATCAAGA 1200
DB |||||||
QY 1141 AGAACTCTAAGGTATATTTGACAAATCAATTCAGAACTTTTAAAAAGAAAGATCAAGA 1200
DB |||||||
QY 1201 GGCAATAGAAAGACAGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAAACAA 1260
DB |||||||
QY 1201 GGCAATAGAAAGACAGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAAACAA 1260
DB |||||||
QY 1261 ACAAGGCTCTAAGACACAGAGGAGGAGAGAGAGAGTGAAGACAGACAGAG 1320
DB |||||||
QY 1261 ACAAGGCTCTAAGACACAGAGGAGGAGAGAGAGAGTGAAGACAGACAGAG 1320
DB |||||||
QY 1321 AAAAAACAGACAGAGAGACAGACAGACAGAGAGAGGCGGAGAGGATAGAAGA 1380
DB |||||||
QY 1321 AAAAAACAGACAGAGAGACAGACAGACAGAGAGAGGCGGAGAGGATAGAAGA 1380
DB |||||||
QY 1381 GAGAGGGGTGGAGAGACACAGATATTTGAGAGAGACTCAGAAAGATAGCCGGGA 1440
DB |||||||
QY 1381 GAGAGGGGTGGAGAGACACAGATATTTGAGAGAGACTCAGAAAGATAGCCGGGA 1440
DB |||||||
QY 1441 GAACACAGAGATGGAAGAGACTCTGAGAAAAAACCCAGAGCAAAAGATGGAAGAGG 1500
DB |||||||
QY 1441 GAACACAGAGATGGAAGAGACTCTGAGAAAAAACCCAGAGCAAAAGATGGAAGAGG 1500
DB |||||||
QY 1501 AGTATCAGGGTGAACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560
DB |||||||
QY 1501 AGTATCAGGGTGAACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560
DB |||||||
QY 1561 TCCAGGCGCCAAAGATAGTGAACCCAGAGTGGTGAAGAGCCAGATCCCTTAAGGCTGGGG 1620
DB |||||||
QY 1561 TCCAGGCGCCAAAGATAGTGAACCCAGAGTGGTGAAGAGCCAGATCCCTTAAGGCTGGGG 1620
DB |||||||
QY 1621 AGGCAGGAAAGGGCTGGCTGGCTCCGGAGACCCCTCCCAATCTCCGGGCGAGGAG 1680
DB |||||||
QY 1681 GTAGGGATGACATTCGGACTGGGTGGGGGTCTCTGGGGTGGAGATAGGGGAGCA 1740
DB |||||||
QY 1681 GTAGGGATGACATTCGGACTGGGTGGGGGTCTCTGGGGTGGAGATAGGGGAGCA 1740
DB |||||||
QY 1741 GGAGGACTATTTGCTAAGGCCCGATAGGACCTCATTTGCCCGGGAATGTGCCCCAGGGAG 1800
DB |||||||
QY 1741 GGAGGACTATTTGCTAAGGCCCGATAGGACCTCATTTGCCCGGGAATGTGCCCCAGGGAG 1800
DB |||||||

QY 1801 CAGTGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCAGAGGAGGAGGAGTGCACAGGA 1860
DB |||||||
QY 1801 CAGTGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCAGAGGAGGAGTGCACAGGA 1860
DB |||||||
QY 1861 AGGCACAGGCTTGAGAGTCTGCGGCTGAGCTGGAGCAAAATCCCCACCCCTACCTGG 1920
DB |||||||
QY 1861 AGGCACAGGCTTGAGAGTCTGCGGCTGAGCTGGAGCAAAATCCCCACCCCTACCTGG 1920
DB |||||||
QY 1921 GGCACAGGCAAGTGAACCTGTGAGGTGGCTCAGCAGCGAGGAGGAGAGGTGTCT 1980
DB |||||||
QY 1921 GGCACAGGCAAGTGAACCTGTGAGGTGGCTCAGCAGCGAGGAGGAGAGGTGTCT 1980
DB |||||||
QY 1981 GTGCTCTCAGCCACCATCTTCTCTGTCCTCTCTGTCCTCTGAGGCTGTCTGAGGCTGTAGA 2040
DB |||||||
QY 1981 GTGCTCTCAGCCACCATCTTCTCTGTCCTCTCTGTCCTCTGAGGCTGTCTGAGGCTGTAGA 2040
DB |||||||
QY 2041 CTCTATCTTTCTGAATTTCTATAGTCTGGGTCTCAGCGCAGTGCAGATGGTCCCGTCTC 2100
DB |||||||
QY 2041 CTCTATCTTTCTGAATTTCTATAGTCTGGGTCTCAGCGCAGTGCAGATGGTCCCGTCTC 2100
DB |||||||
QY 2101 CTTGTGGTTCTCTCTACCTGGGGAATAAGGTAGGAGGAGGAGGAGTGGGTAAAG 2160
DB |||||||
QY 2101 CTTGTGGTTCTCTCTACCTGGGGAATAAGGTAGGAGGAGGAGGAGTGGGTAAAG 2160
DB |||||||
QY 2161 GCTCCCGGATCCGCTGGGCTCCCAACCTCTGACATTTCCCATCCAGGTGCAGCGGCC 2220
DB |||||||
QY 2161 GCTCCCGGATCCGCTGGGCTCCCAACCTCTGACATTTCCCATCCAGGTGCAGCGGCC 2220
DB |||||||
QY 2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGTCTCTGTATCAGAGCTTGCCTT 2280
DB |||||||
QY 2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGTCTCTGTATCAGAGCTTGCCTT 2280
DB |||||||
QY 2281 CTGGGGTCAACAGGTAAACAGAACTCTGGGGTGGAGGGTGTGGGAATGGGAGGACTGT 2340
DB |||||||
QY 2281 CTGGGGTCAACAGGTAAACAGAACTCTGGGGTGGAGGGTGTGGGAATGGGAGGACTGT 2340
DB |||||||
QY 2341 CTCTGGGCACTAGAGCGCTGCTCCCTGGGAACTGTGTGAGCCTGGGCACTGACTCCGG 2400
DB |||||||
QY 2341 CTCTGGGCACTAGAGCGCTGCTCCCTGGGAACTGTGTGAGCCTGGGCACTGACTCCGG 2400
DB |||||||
QY 2401 GACCGGTGAATGTGAGTCTCTGTCTGTACTGTGTGGTTGTGCCATCGTATGTGGCCCTGT 2460
DB |||||||
QY 2401 GACCGGTGAATGTGAGTCTCTGTCTGTACTGTGTGGTTGTGCCATCGTATGTGGCCCTGT 2460
DB |||||||
QY 2461 GACTGCCACGGTGTGTCTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCGG 2520
DB |||||||
QY 2461 GACTGCCACGGTGTGTCTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCGG 2520
DB |||||||
QY 2521 CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTTTGTGATTTGTGCTGCAATTTAAG 2580
DB |||||||
QY 2521 CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTTTGTGATTTGTGCTGCAATTTAAG 2580
DB |||||||
QY 2581 ATTGTGTGGCTCCACAGCTGTGTGGTCAATGCAATGACACTGGGGGTGTTCACTGT 2640
DB |||||||
QY 2581 ATTGTGTGGCTCCACAGCTGTGTGGTCAATGCAATGACACTGGGGGTGTTCACTGT 2640
DB |||||||
QY 2641 GTGTTTGGCTGTGTGGTGAATTTGGCAATTTATATGACTGCAGGTATCTGCAGTTCCTG 2700
DB |||||||
QY 2641 GTGTTTGGCTGTGTGGTGAATTTGGCAATTTATATGACTGCAGGTATCTGCAGTTCCTG 2700
DB |||||||
QY 2701 TCCCTGAGGTCGGGATTCGGTGCACAAAGTGTGTATCACCATGGAAAGCTGTGACT 2760
DB |||||||
QY 2701 TCCCTGAGGTCGGGATTCGGTGCACAAAGTGTGTATCACCATGGAAAGCTGTGACT 2760
DB |||||||
QY 2761 GTGTGCTGTTCAGCGGATTTATGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2820
DB |||||||
QY 2761 GTGTGCTGTTCAGCGGATTTATGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2820
DB |||||||
QY 2821 TTTGTGACCCGTGTGACTTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG 2880
DB |||||||
QY 2821 TTTGTGACCCGTGTGACTTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG 2880
DB |||||||
QY 2881 TCTGTGTGAGGCCGTGTAAATGCTATCTGTATGTGTGATGTGAGCTGTGTCTGAGT 2940
DB |||||||

Db	2881		TCTGTGTGAGCCGCTGTAAATGCTACTGTATGTGTGATGTGTCAGCTGTGTGTCTGGAGT	2940
Qy	2941		TTCTGTCTCTCCCTGGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGGAGATGGGTGC	3000
Db	2941		TTCTGTCTCTCCCTGGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGGAGATGGGTGC	3000
Qy	3001		CAGGTGACTGACTTGTCAGATGTGTGCCCTGTGTGCAGAAAGATATGTGCACTGTGAACATC	3060
Db	3001		CAGGTGACTGACTTGTCAGATGTGTGCCCTGTGTGCAGAAAGATATGTGCACTGTGAACATC	3060
Qy	3061		TGTGCACACAGGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGGTGTGCGATCCC	3120
Db	3061		TGTGCACACAGGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGGTGTGCGATCCC	3120
Qy	3121		GCTAGGCTGCCCGGAGCGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACCTGTG	3180
Db	3121		GCTAGGCTGCCCGGAGCGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACCTGTG	3180
Qy	3181		GAGCAACATGGGCGTCTCTGCAGAACTGCGTGCCTTTGGCTGTTACTGCTGTTGTGC	3240
Db	3181		GAGCAACATGGGCGTCTCTGCAGAACTGCGTGCCTTTGGCTGTTACTGCTGTTGTGC	3240
Qy	3241		CGCTGGTTCCTTGGGTCAGTTCTGTAATGATGTGTGCCAGGCCATCAGCAAGGGTAA	3300
Db	3241		CGCTGGTTCCTTGGGTCAGTTCTGTAATGATGTGTGCCAGGCCATCAGCAAGGGTAA	3300
Qy	3301		GAACAGGCGGGCGGCGTGCCTCACGCCCTGTAATCCACGCCCTTTGGGAGGCCGAGGCA	3360
Db	3301		GAACAGGCGGGCGGCGGCTGCCTCACGCCCTGTAATCCACGCCCTTTGGGAGGCCGAGGCA	3360
Qy	3361		GGCGGATCACCCTGAGGTCGGGAGATCGAGGCCAGCCTGACCAATGGAGAACTCCAGCTACTC	3420
Db	3361		GGCGGATCACCCTGAGGTCGGGAGATCGAGGCCAGCCTGACCAATGGAGAACTCCAGCTACTC	3420
Qy	3421		CTACTAAAAATACAAAAAATTAGCTGTGTGTGTGGCGCGTCCCTGTAAATCCAGCTACTC	3480
Db	3421		CTACTAAAAATACAAAAAATTAGCTGTGTGTGTGGCGCGTCCCTGTAAATCCAGCTACTC	3480
Qy	3481		GGGAGACTGGGCAGAAAAATCCCTTGAAACCCGGGAGGTGAGGTTCCGCTGAGCCGAGA	3540
Db	3481		GGGAGACTGGGCAGAAAAATCCCTTGAAACCCGGGAGGTGAGGTTCCGCTGAGCCGAGA	3540
Qy	3541		TCGCGCCATTGCACTCCAGCCTGGGCAACAGAGCGAAAACTCCGCTCTCGAAAAAATAAAA	3600
Db	3541		TCGCGCCATTGCACTCCAGCCTGGGCAACAGAGCGAAAACTCCGCTCTCGAAAAAATAAAA	3600
Qy	3601		GAATAAAAAAGGGTAAGAACCAAGTGAATGGGCAACCGGAGGACTGATGATGGAGTGGGGC	3660
Db	3601		GAATAAAAAAGGGTAAGAACCAAGTGAATGGGCAACCGGAGGACTGATGATGGAGTGGGGC	3660
Qy	3661		ATGCATCTAGCTGTGTGTGTGTGAGAGGAGGATTTGACAGGATTTGAGAGGCGAT	3720
Db	3661		ATGCATCTAGCTGTGTGTGTGTGAGAGGAGGATTTGACAGGATTTGAGAGGCGAT	3720
Qy	3721		GTTTTTCATCTGAGAAATTCAGAAACCTAGGCGCTGCTTTCCCGCTCCATGTGGCCCCCTAAG	3780
Db	3721		GTTTTTCATCTGAGAAATTCAGAAACCTAGGCGCTGCTTTCCCGCTCCATGTGGCCCCCTAAG	3780
Qy	3781		CTGAGCCCTTCTTTCTGTCGTCCTGTTTCGAAACCCCTAGCTCCGCCCATGAGCTCTGACC	3840
Db	3781		CTGAGCCCTTCTTTCTGTCGTCCTGTTTCGAAACCCCTAGCTCCGCCCATGAGCTCTGACC	3840
Qy	3841		CCACCTCCTTTCTCAACACAGCCCCCTAGGCCAGACTCTAGTGGACCCCGCCTTAAGGCCA	3900
Db	3841		CCACCTCCTTTCTCAACACAGCCCCCTAGGCCAGACTCTAGTGGACCCCGCCTTAAGGCCA	3900
Qy	3901		CACCCCTTTGGGCCAGGCTCACCCCTTATCTGTGGGTACCTTTCTAGAACCCCTTTCAA	3960
Db	3901		CACCCCTTTGGGCCAGGCTCACCCCTTATCTGTGGGTACCTTTCTAGAACCCCTTTCAA	3960
Qy	3961		AGTCAGAGCTTTTCTTTTCTTTTGGAGACAGTCTTGCTCTCTCTCCAGGCTGGAG	4020

Db	3961	AGTCAGAGCTTTT	TTTTTTTTTTGGAGACGTCTTGCTCTCTCTCCAGCGTGGAG	4020
Qy	4021	TGCAGTGGCGTGATCTCGGCTCACTGCACCAACCTCTGCCTCCAGGTTCAAGTGATCTCGT		4080
Db	4021	TGCAGTGGCGTGATCTCGGCTCACTGCACCAACCTCTGCCTCCAGGTTCAAGTGATCTCGT		4080
Qy	4081	GCCTCCACCTCTCGAGTAGCTGGGATTAACAGGTGCGCGCCAACAAGCTGGCTGATTTTTT		4140
Db	4081	GCCTCCACCTCTCGAGTAGCTGGGATTAACAGGTGCGCGCCAACAAGCTGGCTGATTTTTT		4140
Qy	4141	GTGCTTTTAGTAGAGACAGGGTTTACCTTGTGTGGCAGGCTGGTCTCAAACTCCCAACC		4200
Db	4141	GTGCTTTTAGTAGAGACAGGGTTTACCTTGTGTGGCAGGCTGGTCTCAAACTCCCAACC		4200
Qy	4201	TCAGTGATCCGCGCACCTCGGCTCCAGAGTGTCTGGGTTTACAGGCGTGAGCCACCGC		4260
Db	4201	TCAGTGATCCGCGCACCTCGGCTCCAGAGTGTCTGGGTTTACAGGCGTGAGCCACCGC		4260
Qy	4261	CCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTTAACATGTAAACCTCGACCTGGCC		4320
Db	4261	CCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTTAACATGTAAACCTCGACCTGGCC		4320
Qy	4321	CTAACTAAGTCAATTCCAAAACCTTCTGCTCCAGGCCCTGACGCCACTCACTGAGGCC		4380
Db	4321	CTAACTAAGTCAATTCCAAAACCTTCTGCTCCAGGCCCTGACGCCACTCACTGAGGCC		4380
Qy	4381	TGACCCCACTTCTTGAGACCAAGTCCATCCCTTAAAGCCCTGTGCTCCCATCCCAAT		4440
Db	4381	TGACCCCACTTCTTGAGACCAAGTCCATCCCTTAAAGCCCTGTGCTCCCATCCCAAT		4440
Qy	4441	GCTCCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAAT		4500
Db	4441	GCTCCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAAT		4500
Qy	4501	TTTACCTTCAATGTAGTTCTAGCCAAATCCAGGAATCTGTGAGGTCCAGTTAGTCCA		4560
Db	4501	TTTACCTTCAATGTAGTTCTAGCCAAATCCAGGAATCTGTGAGGTCCAGTTAGTCCA		4560
Qy	4561	GTAACCTTACCTGAGCGTGGGCTCTGCTTGTAGCTTGTGAGCTTGGGCTTGGAGGTGCCA		4620
Db	4561	GTAACCTTACCTGAGCGTGGGCTCTGCTTGTAGCTTGTGAGCTTGGGCTTGGAGGTGCCA		4620
Qy	4621	CTCTTATTTCTCCAGGCCCTGCCCCTGCCCCCTCAGCATGTTCAGACACCCACCTCTAGCT		4680
Db	4621	CTCTTATTTCTCCAGGCCCTGCCCCTGCCCCCTCAGCATGTTCAGACACCCACCTCTAGCT		4680
Qy	4681	GGTCTGGCTTGTAGTCTGAAACCCACCCAGCCCCAAGCCCGGCTCTGAGCCCCGCC		4740
Db	4681	GGTCTGGCTTGTAGTCTGAAACCCACCCAGCCCCAAGCCCGGCTCTGAGCCCCGCC		4740
Qy	4741	CAACCCATTTTCCGTTCCAGAGCATGTTCTGCCCCAACAAATGATTTCTGTGACACC		4800
Db	4741	CAACCCATTTTCCGTTCCAGAGCATGTTCTGCCCCAACAAATGATTTCTGTGACACC		4800
Qy	4801	CCTCTAACACCGTCCCTCTGGGAGCAACAGGAGCTGGGAGCTGGGGCCGGGGAAGACG		4860
Db	4801	CCTCTAACACCGTCCCTCTGGGAGCAACAGGAGCTGGGAGCTGGGGCCGGGGAAGACG		4860
Qy	4861	CCCGGTGGATGACAGCAGCAGCGCATCATATGGATCCGACTCCGATATGCAACACC		4920
Db	4861	CCCGGTGGATGACAGCAGCAGCGCATCATATGGATCCGACTCCGATATGCAACACC		4920
Qy	4921	AGCCGTGGCAGGCCGCGTGTGCTAAGGCCCAACAGCTCTACTCGGGGCGGTGTGG		4980
Db	4921	AGCCGTGGCAGGCCGCGTGTGCTAAGGCCCAACAGCTCTACTCGGGGCGGTGTGG		4980
Qy	4981	TGCATCCACAGTGGCTGCTC	5000	
Db	4981	TGCATCCACAGTGGCTGCTC	5000	

RESULT 5
ACH80345

ACH80345 standard; DNA; 525 BP.
ACH80345;
29-JUL-2004 (first entry)
Human genome derived single exon probe #13540.
Human; probe; ss; gene expression; single exon probe; microarray;
alternative splicing event; genomic alteration.
Homo sapiens.
US2003194704-A1.
16-OCT-2003.
03-APR-2002; 2002US-00029386.
03-APR-2002; 2002US-00029386.
(PENW/) PENN S G.
(RANK/) RANK D R.
(HANZ/) HANZEL D K.
Penn SG, Rank DR, Hanzel DK;
WPI; 2004-119264/12.
New human genome-derived single exon nucleic acid probes useful for human
gene expression analysis, for identifying or characterizing alternative
splicing events, for assessing genomic alterations or as tools for
surveying tissues.
Claim 15; SEQ ID NO 13540; 80pp; English.
The invention relates to a nucleic acid probe for measuring human gene
expression, comprising any of the 27,400 fully defined nucleotide
sequences in the specification, or their complements or fragments, and
encoding at least 8 amino acids of any of the 6888 amino acid sequences
fully defined in the specification. The probe is a single exon probe that
hybridises under high stringency conditions to a nucleic acid molecule
expressed in human cells or tissues. Also included are a spatially-
addressable set of single exon nucleic acid probes for measuring human
gene expression (comprising a plurality of single exon nucleic acid
probes cited above, where each of the plurality of probes is separately
and addressably isolatable or amplifiable from the plurality), a single
exon microarray for measuring human gene expression, a method of
measuring human gene expression, a vector comprising the single exon
probe cited above, an ORF-encoded peptide comprising at least 8
contiguous amino acids of any of the above-mentioned amino acid
sequences (optionally with conservative amino acid substitutions), an
isolated antibody that binds specifically to a peptide cited above,
methods of selling and/or licensing single exon probes or microarrays to
a customer desiring to measure gene expression, a method of providing
human gene expression data by subscription, and a computer-readable
storage medium which contains a database having a plurality of records
(each record including data on the expression of a single exon probe
cited above). The probe, methods and apparatus are useful in gene
expression analysis. The probes may be used as tools for surveying
tissues to detect the presence of expressed messages that contain their
specific exon, or in constructing genome-derived single exon microarrays.
In addition, the probes are used in identifying and characterising
alternative splicing events, in detecting and characterising gross
alterations in the genomic locus that includes their exon, in assessing
smaller genomic alterations, in priming the synthesis of nucleic acids,
or in expressing the ORF-encoded peptide. The present sequence is a human
single exon probe of the invention. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20030194704
Sequence 525 BP; 99 A; 178 C; 145 G; 103 T; 0 U; 0 Other;

Query Match 9.4%; Score 472; DB 12; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.3e-76;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4529 TCCAGGAATCTGTGAGGTCAGTTAGAGTCCAGTAACCCCTACCTGAGCCCTGGGCTCTGTC 4588
Db 1 TCCAGGAATCTGTGAGGTCAGTTAGAGTCCAGTAACCCCTACCTGAGCCCTGGGCTCTGTC 60
Qy 4589 CTTGAGCTTGAGCCTGGGCTTGAGAGTGCCACTCTTATTCTCCAGGCCCTGCCCTGCC 4648
Db 61 CTTGAGCTTGAGCCTGGGCTTGAGAGTGCCACTCTTATTCTCCAGGCCCTGCCCTGCC 120
Qy 4649 CCCTCAGCATGTGAGACACCCCTCTAGTGTGTGGCTCTTGTAGTCTGAAACCCAC 4708
Db 121 CCCTCAGCATGTGAGACACCCCTCTAGTGTGTGGCTCTTGTAGTCTGAAACCCAC 180
Qy 4709 CCCAGGCCCAAGCCCGCCCTCTGAGCCCGCCCAACCCATTTTCGTTCCAGAGCATGT 4768
Db 181 CCCAGGCCCAAGCCCGCCCTCTGAGCCCGCCCAACCCATTTTCGTTCCAGAGCATGT 240
Qy 4769 TCTCGCCCAACAATGATGTTTCTGTGTGACACCCCTCTTAAACCGTGCCTCTGGAGCAA 4828
Db 241 TCTCGCCCAACAATGATGTTTCTGTGTGACACCCCTCTTAAACCGTGCCTCTGGAGCAA 300
Qy 4829 CCAGGACCTGGAGCTGGGCGCGGGAAGACCGCGTCCGATGACAGCAGCAGCCGAT 4888
Db 301 CCAGGACCTGGAGCTGGGCGCGGGAAGACCGCGTCCGATGACAGCAGCAGCCGAT 360
Qy 4889 CATCAATGGATCCGACTGCGATATGCACACCCAGCCCGCTGGCAGCCCGCTGTTGCTAAG 4948
Db 361 CATCAATGGATCCGACTGCGATATGCACACCCAGCCCGCTGGCAGCCCGCTGTTGCTAAG 420
Qy 4949 GCCCAACACAGCTTACTGCGGGCGGTGTTGGTGATCCACAGTGGTGTCTC 5000
Db 421 GCCCAACACAGCTTACTGCGGGCGGTGTTGGTGATCCACAGTGGTGTCTC 472

RESULT 6
ACN44942
ID ACN44942 standard; DNA; 41454 BP.
XX AC ACN44942;
XX AC ACN44942;
DT 18-NOV-2004 (first entry)
XX DT
DE Human genomic sequence hCG21040.
XX DE
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX KW
OS Homo sapiens.
XX OS
PN WO2003073826-A2.
XX PN
PD 12-SEP-2003.
XX PD
PF 28-FEB-2003; 2003WO-US006235.
XX PF
PR 01-MAR-2002; 2002US-00087192.
XX PR
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW;
PI
XX WPI; 2003-328604/31.
DR
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 1642; Opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for

Qy	314	CACAGTTTTCAGGGTGTCTAATGATCACTCTTTT	373.
Db	20797	CACAGTGCCCTATTTCCTCTTCCAAGTATGATTCCTTTT	20738
Qy	374	TTGAGACGGAGTCTCGCTCTGTCCGCCAGCGCTGGAGTGCAGTGGCGGGATCTCGGCTCAC	433
Db	20737	TTGAGACGGAGTCTCGCTCTGTCTGTCCGCCAGCGCTGGAGGCGAGTGGCGGNATCTTGCGCTCAC	20678
Qy	434	TGCNAGCTCCGCTCCGGGTTCACGCCATTTCTCTGCTCAGCTCCCAAGTAGCTGGG	493
Db	20677	TGAAGCTCCGCTCCGGGTTCACGCCATTTCTCTGCTCAGCTCCCGAGTAGCTGGG	20618
Qy	494	ACTACAGGCCCGCCCACTACGCCCGGCTAAATTTTTTCTGTAATTTTTTAGTAGAGACGGGT	553
Db	20617	ACGACAGGCCCGCCCACTACGCCCGGCTAA- TTTTITGTAATTTTAGTAGAGACGGGT	20559
Qy	554	TTCAACGTTTTAGCGGGATGGCTCGATCTCTGACCTCGTGATCGCGCCGCTCGGCC	613
Db	20558	TTCAACGTTTTAGCCAGGATGGTCTCGATCTCTGACCTCGTGATCGGCCACCTCGGCC	20499
Qy	614	TCCCAAAGTCTGGGATTACAGSCGTGAGCCACCGCGCCGCCCATGATCATCT-----	667
Db	20498	TCCCAAAGTCTGGGATTACAGSGTGTAGCCACCGCGCCGCCCGTATGATTTCTTAAA	20439
Qy	668	-----TCTTGACTATGCTGATGCACAAGTAC	694
Db	20438	TGTGCTTATTTTGGTTTAACTCTTATTTTGAGTTTTTTGGATACAACTAGAACAAGCAC	20379
Qy	695	CTAAGCCATCA-GACTCTACCTTTAAATATGCAAGTTTGGGCGCAGGCACCGTGCTCAT	753
Db	20378	AATAAATAATCATCTTATCCACCGTCTAAATACAAATTTGTAGGCCAGGCACCGTGACTCAC	20319
Qy	754	GCTGTAAATCCAGCACTTTGGGAGCGAGAGTGGGTGAATCACTTTGAGGCCAGAGTTTT	813
Db	20318	ACCTGTAAATCCCAACACTTTGAGAGGCCGAGATGGSCAACTCAC--GAGTCAAGGAGTTC	20261
Qy	814	GAGACCAGCTGGCCAACATGTTGAACHTCTGCTTTACTAAAAAAMAAAAAAAAA	873
Db	20260	GAGACCAGCTGGCCAACATGTTGAACCCCACTTATT-----AAAAATAC	20213
Qy	874	AAAAATCAGCCGGGTGTGTTGGGGCACACTGTAATCCCAGACTATGCTGGAGGCTGAGGC	933
Db	20212	AAAAATTAGCTGGGTGTGGTGGCCAGGCTGTATCCAGACTCTCGGAGGCTGAGGC	20153
Qy	934	ACGAGATGACTTTGAACCTCTGAGGCGGAGGTTGCAGTGGGCGGAGATCAACATCACGCC	993
Db	20152	ATGAGAAATTTGCTTTAGCCCCGGGAGGAGAGTTTGCAGTGAGCTGAGAACATGCCATTGTA	20093
Qy	994	CTCCAGCTGGGGCAGACAGCAAGACTCTGTCTCAAAATAAATAAATAACAAACGACAA	1053
Db	20092	GTCACGCTGGGCACAGACGAGACTCTCTTCTCAATAAATAAATAAATAAATAAAGTT	20033
Qy	1054	GCAGTTTGTGTACCTTAGTTATATCAAAAAAAT	1091
Db	20032	GTTGATCTTTCACCATATCATTTAACTGATAAAGAT	19995

RESULT 8

RESULI 8
ADN94799

ADN54733
ID ADN94799 standard: DNA: 91352 BP.

XX XX

ADN
AC

XX

DT 12-

XX

DE DNA

XX

KW gen.

KW
Chem

XX

SO
HOM

XX

PN	US2004097451-A1.
XX	
PD	20-MAY-2004.
XX	
PF	19-NOV-2002; 2002US-00300611.
XX	
PR	19-NOV-2002; 2002US-00300611.
XX	
PA	(ISIS-) ISIS PHARM INC.
XX	
PI	Chiang M, Dobie KW;
XX	
DR	WPI; 2004-389192/36.
XX	
PT	New compounds, particularly oligonucleotides targeted to a nucleic acid
PT	encoding nidogen, useful for treating diseases associated with nidogen,
PT	e.g. Chediak-Higashi syndrome.
XX	
PS	Claim 1; SEQ ID NO 4; 91pp; English.
XX	
CC	The invention relates to antisense oligonucleotides which are targeted
CC	to, and inhibit the expression of, a nucleic acid molecule encoding
CC	nidogen. The antisense oligonucleotides are useful for treating a disease
CC	or condition associated with nidogen, such as Chediak-Higashi syndrome.
CC	They are also useful in research and diagnostics for modulating the
CC	expression of nidogen. The present sequence represents the gene encoding
CC	human nidogen.
XX	
SQ	Sequence 91352 BP; 23543 A; 20675 C; 20522 G; 26612 T; 0 U; 0 Other;
	Query Match 8.4%; Score 419; DB 12; Length 91352;
	Best Local Similarity 79.4%; Pred. No. 3e-66;
	Matches 552; Conservative 0; Mismatches 125; Indels 18; Gaps 4;
Qy	350 TTTTCTTTTTTTTTTTTTTTTTTTTTTGAGCGAGCTCTGCTCTGTGCCCAGGCTGGAG 409
Db	11334 TTATTATTAATAATATATATATTTTGAAGTAGAGTCTTGCTCTGTGCCCAGGTTGGAG 11393
Qy	410 TGCA GTGGCGGATCTCGGCTCACATGC AAGTC CGCCTCCCGGGTTACG GCATTCTCCT 469
Db	11394 TGCA GTGGTGTGCTCTTG GCTTACTGCA AGCTCCACCTCCG GGTTC CGGCATCTCCT 11453
Qy	470 GCCTCAGCTCCC AAGTAGCTGGGACTACAGCGCCCGCCACTACGCGCGGCTAAATTTT 529
Db	11454 GCCTCAGCTCCCGAGTAGCTGGGACTACAGACACCGCCACCAGCCCGGCTAA-TTTT 11512
Qy	530 TTGTATTTTAGTAGAGCGGGTTTACCGTTT TAGCGGGATGGCTCATCTCTCTGA 589
Db	11513 TTGTATTTTAGTAGAGCGGGTTTACCGGTTT ACCCGGTTTACCAGGATGGTCTCGATCTCTGA 11572
Qy	590 CCTCGTGATCCGCGCCCTCGGCCTCCCAAAGTGTGGGATTCAGGCGGTGAGCCACCGC 649
Db	11573 CCTCGTGATCCACCCTCGGCCTCCCAAAGTGTGGGATTCAGGCGGTGAGCCACCGC 11632
Qy	650 GCCCGGCCATGATCATCTTCTTGACTATGC----TGATGTGACAAGTACCTAAAAGCCATC 705
Db	11633 AC CGGCCCTGAATTATTTTAAGCAATCCAATATTTTATCGTTGTGAATGTTT 11692
Qy	706 AGACTCTACCTTTAAATATG-CAGTTTGGCCAGGACCGGTGGCTCATGCTGTAAATTC 764
Db	11693 AGATGTGCTTCAGAAGAAAGACTCTTTAGGCCGAGCAGGTGGGCTCATACCTGTAAATTC 11752
Qy	765 CAGCACTTTGGGAGCAGAGGTGGGTGAATCATTTGAGCCAGGAGTTTGAGACCAAGCT 824
Db	11753 CAGCACTTTGGGAGGCTGAGCGGGCGGATAGCTTTGAGCTCAGGAGTTTCCAGATAGCCT 11812
Qy	825 GGCCAACATGGTGAACCTCTGTCTTTACTAATAAAAAAATAAAAAAAAAAATCAGCC 884
Db	11813 GGCCAACATGGTGAACCTCTGTCTCACT-----AATAATATCAAAATAGCC 11860
Qy	885 GGGTGTCTGGGGCAACCTGTAATCCAGCTATGCTCGAGGCTGAGCACGAGAGTCAAC 944
Db	11861 AGATGTGGTGGCACACACTGTGATCTCCAGCTACTCGGGAGGCTGAGCAGGAAATGC 11920

KW	Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.	
XX	Homo sapiens.	
OS	WO2003073826-A2.	
XX	12-SEP-2003.	
XX	28-FEB-2003; 2003WO-US006235.	
XX	01-MAR-2002; 2002US-00087192.	
XX	(SAGR-) SAGRES DISCOVERY.	
XX	Morris DW;	
XX	WPI; 2003-328604/31.	
DR	Recombinant nucleic acid useful for diagnosis and treatment of carcinoma	
XX	comprises a nucleotide sequence.	
PT	Claim 1; SEQ ID NO 1450; Opp; English.	
XX	The present invention relates to novel DNA and protein sequences which	
CC	are associated with carcinomas. The sequences are useful for: (i) for	
CC	screening drug candidates; (ii) for screening of bioactive agent capable	
CC	of binding to Carcinoma Associated Protein (CAP); (iii) for screening of	
CC	a bioactive agent capable of modulating the activity of CAP; (iv) for	
CC	evaluating the effect of a candidate carcinoma drug; (v) for diagnosing	
CC	carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating	
CC	carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;	
CC	(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for	
CC	determining Carcinoma Associated (CA) gene copy number. In addition, the	
CC	CA genes are useful as DNA vaccines and the CAP are useful as markers of	
CC	carcinoma including lymphoma. The present sequence is one such CA coding	
CC	sequence. Note: This patent is an equivalent to basic patent	
CC	US2002182586A1, for which no sequence data was published	
XX	Sequence 26371 BP; 6652 A; 6411 C; 6495 G; 6813 T; 0 U; 0 Other;	
XX	Query Match 8.4%; Score 417.6; DB 11; Length 26371;	
XX	Best Local Similarity 78.0%; Pred. No. 4.7e-66;	
XX	Matches 551; Conservative 0; Mismatches 129; Indels 26; Gaps 3;	
QY	355 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	414
DB	4590 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	4649
QY	415 TGGCGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGTTTCAGCCATTCTCTGCTC	474
DB	4650 TGGCGGATCTTGGCTCACTGCAAGCTCCACCTCCCGGTTTCAGCCATTCTCTGCTC	4709
QY	475 AGCTCCCAAGTAGCTGGGACTACAGCGCGCGCTACGCGCGCTAAATTTTGTGTA	534
DB	4710 AGCTCCCAAGTAGCTGGGACTACAGCGGTTGCGCACACGCTGGGTAA-TTTTGTGTA	4768
QY	535 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	594
DB	4769 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	4828
QY	595 TGATCCCGCCCTCGGCTCCCAAGTGTGGGATTACAGGCTGAGCCACCGCGCCG	654
DB	4829 CGATCCCGCCCACTCGGCTCCCAAGTGTGGGATTACAGGCTGAGCCACCGCGCCG	4888
QY	655 GCGATGATCATCTCTTTGACTATGCTGATGTGCAAGTACCTAAAGCCATCAGACTCTAC	714
DB	4889 ACCCGAGTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT	4935
QY	715 CCTTTAAATATGCAAGTTTGGGCCAGGACCGCTGCTCATGCTGTAATTCAGCACTTTG	774
DB	4936 ATTTAAAGAGAGAAACTGGCGGGTGGCTGCTCAGCGCTGTAATCCAGCACTTTG	4995
QY	775 GGAGGACAGAGTGGTGAATCACTTTAGGAGGAGAGTTTGAACCCAGCTGGCCCAACATG	834

DB	4996 CGAGACCCAGCGGCGAGATCATTGAGGTCAGAGTTTGAGATAGCTTGGCCCAATG	5055
QY	835 GTGAACTCTGTCTTTACTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	894
DB	5056 GTGAAACCCCATCTCTACT-----AAAATACAAAATTTAGCCGGCATGGTG	5103
QY	895 GGGCACACCTGTAATCCAGCTATGCTGAGGCTGAGGCGAGGACGAGTCACTTGAACCTG	954
DB	5104 GCGGCGCCTGTAATCCAGCTACTCGGAGGCTGAGGCGAGGAGATAGCTTGAACCCGG	5163
QY	955 GAGCGGAGGTTGAGTGGGCGGAGATCACATCACCGCCCTCCAGCCCTGGGCGGACAGC	1014
DB	5164 GAGCGGAGGTTGAGTGGGCGGAGATCACATCACCGCCCTCCAGCCCTGGGCGGAGT	5223
QY	1015 AAGACTCTGTCTCAATAATAATAAACAACAAACGACAGCAGTTT	1060
DB	5224 GAGACTCTGTCTCAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATA	5269
XX	RESULT 11	
XX	ACN44410	
ID	ACN44410 standard; DNA; 91760 BP.	
XX	ACN44410;	
XX	18-NOV-2004 (first entry)	
DE	Human genomic sequence hCG41365.	
KW	Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.	
OS	Homo sapiens.	
PN	WO2003073826-A2.	
XX	12-SEP-2003.	
XX	28-FEB-2003; 2003WO-US006235.	
XX	01-MAR-2002; 2002US-00087192.	
XX	(SAGR-) SAGRES DISCOVERY.	
XX	Morris DW;	
XX	WPI; 2003-328604/31.	
PT	Recombinant nucleic acid useful for diagnosis and treatment of carcinoma	
XX	comprises a nucleotide sequence.	
XX	Claim 1; SEQ ID NO 844; Opp; English.	
XX	The present invention relates to novel DNA and protein sequences which	
CC	are associated with carcinomas. The sequences are useful for: (i) for	
CC	screening drug candidates; (ii) for screening of bioactive agent capable	
CC	of binding to Carcinoma Associated Protein (CAP); (iii) for screening of	
CC	a bioactive agent capable of modulating the activity of CAP; (iv) for	
CC	evaluating the effect of a candidate carcinoma drug; (v) for diagnosing	
CC	carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating	
CC	carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;	
CC	(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for	
CC	determining Carcinoma Associated (CA) gene copy number. In addition, the	
CC	CA genes are useful as DNA vaccines and the CAP are useful as markers of	
CC	carcinoma including lymphoma. The present sequence is one such CA coding	
CC	sequence. Note: This patent is an equivalent to basic patent	
CC	US2002182586A1, for which no sequence data was published	
XX	Sequence 91760 BP; 21177 A; 23589 C; 24129 G; 22845 T; 0 U; 20 Other;	
XX	Query Match 8.4%; Score 417.6; DB 11; Length 91760;	
XX	Best Local Similarity 76.8%; Pred. No. 5.4e-66;	
XX	Matches 576; Conservative 0; Mismatches 139; Indels 35; Gaps 4;	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:41:10 ; Search time 567.479 Seconds
(without alignments)
14417.061 Million cell updates/sec

Title: US-09-936-271C-13_COPY_1_5000
Perfect score: 5000
Sequence: 1 gggccagactgaagcaag.....tgatccacagtggtgctc 5000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgm2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgm2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgm2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgm2_6/ptodata/1/ina/PTCUS COMB.seq:*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2245	44.9	10818	4	US-09-949-016-13583
C 2	444	8.9	246240	2	US-08-724-394A-20
C 3	444	8.9	246240	2	US-08-724-394A-21
C 4	444	8.9	246240	2	US-08-724-394A-22
5	441	8.8	97376	4	US-09-949-016-16093
6	421.2	8.4	22973	4	US-09-949-016-13644
C 7	417.6	8.4	120727	4	US-09-949-016-15787
C 8	417.6	8.4	120727	4	US-09-949-016-15788
9	415.4	8.3	166698	4	US-09-949-016-16038
10	414.2	8.3	79858	4	US-09-949-016-16080
C 11	414	8.3	60304	4	US-09-949-016-12218
C 12	414	8.3	60305	4	US-09-949-016-15791
C 13	411.2	8.2	125188	4	US-09-949-016-11980
C 14	405.6	8.1	33379	4	US-09-949-016-13861
C 15	404.8	8.1	50217	4	US-09-949-016-16067
C 16	404	8.1	58593	4	US-09-949-016-12232
17	402.6	8.1	53442	4	US-09-949-016-11921
18	402.6	8.1	53453	4	US-09-949-016-13370
C 19	402.2	8.0	64518	4	US-09-949-016-17289
C 20	400.8	8.0	55114	4	US-09-949-016-16792
C 21	400.8	8.0	87617	4	US-09-949-016-16551
C 22	400.2	8.0	81701	4	US-09-949-016-14891
23	400.2	8.0	94855	4	US-09-949-016-12264
24	399.8	8.0	71879	4	US-09-949-016-17465
C 25	399.6	8.0	130563	4	US-09-949-016-12273
C 26	399.6	8.0	131379	4	US-09-949-016-16050
27	399.4	8.0	39920	4	US-09-949-016-16316

C 28	397.2	7.9	79350	4	US-09-949-016-12467	Sequence 12467, A
C 29	397.2	7.9	79351	4	US-09-949-016-16275	Sequence 16275, A
C 30	396.6	7.9	45489	4	US-09-949-016-13398	Sequence 13398, A
C 31	396.6	7.9	87734	4	US-09-949-016-17521	Sequence 17521, A
C 32	396.4	7.9	19451	4	US-09-949-016-13695	Sequence 13695, A
C 33	395.8	7.9	26104	4	US-09-949-016-14045	Sequence 14045, A
C 34	395.8	7.9	77626	4	US-09-949-016-12608	Sequence 12608, A
C 35	395.6	7.9	39920	4	US-09-949-016-16316	Sequence 16316, A
C 36	395	7.9	43255	4	US-09-949-016-11909	Sequence 11909, A
C 37	394.8	7.9	9226	4	US-09-949-016-12596	Sequence 12596, A
C 38	394.8	7.9	9226	4	US-09-949-016-16007	Sequence 16007, A
C 39	394.6	7.9	49487	4	US-09-949-016-15721	Sequence 15721, A
C 40	393	7.9	133613	4	US-09-949-016-15824	Sequence 15824, A
C 41	392.8	7.9	58782	4	US-09-949-016-16851	Sequence 16851, A
C 42	391.6	7.8	69834	4	US-09-949-016-11959	Sequence 11959, A
C 43	391.6	7.8	69834	4	US-09-949-016-12925	Sequence 12925, A
C 44	391.4	7.8	8405	4	US-09-949-016-11882	Sequence 11882, A
C 45	391.4	7.8	8405	4	US-09-949-016-15376	Sequence 15376, A

ALIGNMENTS

RESULT 1

US-09-949-016-13583

; Sequence 13583, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13583

; LENGTH: 10818

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-13583

Query Match 44.9%; Score 2245; DB 4; Length 10818;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2756	TGACTGTGTGCTGCTTCGACGCGATTATGTGATGTGGCTGAGTGTGACGTTATGATGATGC	2815
Db	1	TGACTGTGTGCTGCTTCGACGCGATTATGTGATGTGGCTGAGTGTGACGTTATGATGATGC	60
Qy	2816	CCGTAATTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGACTGT	2875
Db	61	CCGTAATTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGACTGT	120
Qy	2876	GTGTGTCTGTGTGAGGCGGTGTAAATCTACTATGTGTGATGCTGTCAGCTGTGTCT	2935
Db	121	GTGTGTCTGTGTGAGGCGGTGTAAATCTACTATGTGTGATGCTGTCAGCTGTGTCT	180
Qy	2936	GGAGTTTCTGTCTGCTGCGAGGATAGAGGGGTGACGGGTAGCTATCTCTGGGAGATG	2995
Db	181	GGAGTTTCTGTCTGCTGCGAGGATAGAGGGGTGACGGGTAGCTATCTCTGGGAGATG	240
Qy	2996	GGTGCCAGGTGACTGACTTGCAGTGTGTGCTGTGTGTCAGAGAGATGTGGCAGTCTGA	3055
Db	241	GGTGCCAGGTGACTGACTTGCAGTGTGTGCTGTGTGTCAGAGAGATGTGGCAGTCTGA	300
Qy	3056	ACATCTGTGCACACACCGGCATCTGTGCGTGCACACTGTGATGAGGGTGTGCG	3115

	Best Local Similarity	80.1%;	Pred. No. 5.4e-93;		Mismatches	120;	Indels	24;	Gaps	4;
	Matches	580;	Conservative	0;						
Qy	335	TATGATCATCTTC								394
Dd	3310	TTTATATAAGACT								3251
Qy	395	TCGCCCAGGCTGGAGTGCAGTGCGCGGATCTCGGCTCACTGCAAGTCCCGCTCCCGGGT								454
Dd	3250	TCGCCCAGGCTGGAGTGCAGTGCGCGGATCTCGGCTCACTGCAAGTCCCGCTCCCGGGT								3191
Qy	455	TCAGGCCATTCTCTGTGCTCAGCCTCCGAAGTAGCTGGAGCTACAGGCGCCCGCACATC								514
Dd	3190	TCAGCCATTCTCTGTGCTCAGCCTCCGAAGTAGCTGGAGCTACAGGCGCCCGCACATC								3131
Qy	515	GCCCGGCTAAATTTTTTGTATTTTTTACTAGAGACGGGTTTCAACGTTTATAGCCGGGATG								574
Dd	3130	GCCCGGCTAA--TTTTTGTATTTTTTGTAGAGACGGGTTTCAACGTTTATAGCCGGGATG								3072
Qy	575	GCCTCGATCTCTGACCTCGTGTATCCCGCGCCTCGGCCTCCCAAAGTGTGGGATTAACA								634
Dd	3071	GTCTCGATCTCTGACCTCGTGTATCCCGCGCCTCGGCCTCCCAAAGTGTGGGATTAACA								3012
Qy	635	GGCGTAGACCACCGCGCCCGGCCAANGATCATCTTCTTTGACTATGCTGATGTGACAAAGTAC								694
Dd	3011	GGCGTAGACCACCGCGCCCGGCC-----TTATAAAGACTTTTATAGATAAAAAACAGAAT								2959
Qy	695	CTAAGGCCATCAGACTCTACCCCTTAAATATGAGTTTGGGCCAGGCACCGTGGCTCATG								754
Dd	2958	CTATATCTCTAAATGTATG----TATTAAGGATGAACCCCCTGGGCGCAGTGGCTCACA								2903
Qy	755	CCTGTAATCCAGCACCTTTGGGAGCAGAGTGGGTGAATCACTTTGAGGCCACAGAGTTTG								814
Dd	2902	CCTGTAATCCAGCACCTTTGGGAGCCGAGCGGGGTAGATCACTTTGAGTTCAGAGTTTCG								2843
Qy	815	AGACCAGCTGGCCAACATGGTGAACCTGTCTTTTACTAAAAAAAATAAAAAAAAAAAAA								874
Dd	2842	AGACCAGCTGGCCAACACGCGGAAACCCCTGTCTCTACT-----AAAAATACA								2795
Qy	875	AAATACGCGGGTGTCTGTGGGGCACACTGTATAATCCAGCTATGCTGGAGGCTGAGGCA								934
Dd	2794	AAATATAGCTGGGCATGTGGCAGGCACCTGTATAATCCAGCTGTCTTGGGAGACTGAGGCA								2735
Qy	935	CGAGAGTCACTTGAACCCCTGGAGCGGAGGTTCAGTGGGCCGAGATCAATCACCGCCC								994
Dd	2734	GGAAATCGTTTGAACCTGGAGCCTGGGGATFTGAGTSGAGCCAAAGCTGCACCACCTGCAC								2675
Qy	995	TCCAGCCTGGCGCAGACAGCAAGACTCTGTCTCAAAATAAATAAACCAAAACGAACAAG								1054
Dd	2674	TCCAGCCTGGCGCAGACAGCTCTCAAAAAATAAATAAATAAATAAATAAATAAATAA								2615
Qy	1055	CAGT 1058								
Dd	2614	AAAT 2611								

```

RESULT 4
US-08-724-394A-22/c
; Sequence 22, Application US/08724394A
; Patent No. 587237
;
; GENERAL INFORMATION:
;
; APPLICANT: Feder, John N.
; APPLICANT: Kromal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
;
; TITLE OF INVENTION: Megabase Trans-
;
; TITLE OF INVENTION: Sequences and
;
; NUMBER OF SEQUENCES: 31
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: TOWNSEND and TOWNSEND

```

```

, STREET: Two Embarcadero Center, 8th Floor
, CITY: San Francisco
, STATE: CA
, COUNTRY: USA
, ZIP: 94111-3834
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/724,394A
, FILING DATE: 01-OCT-1996
, CLASSIFICATION: 536
, ATTORNEY/AGENT INFORMATION:
, NAME: Fitts, Renee A.
, REGISTRATION NUMBER: 35,136
, REFERENCE/DOCKET NUMBER: 017957-000100
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 415-576-0200
, TELEFAX: 415-576-0300
, INFORMATION FOR SEQ ID NO: 22:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 246240 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: not relevant
, TOPOLOGY: not relevant
, MOLECULE TYPE: cdna
, FEATURE:
, NAME/KEY: misc.feature
, LOCATION: 1..246240
, OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

```

Query Match	8.9%	Score 444;	DB 2;	Length 246240;
Best Local Similarity	80.1%	Pred. No. 5.4e-93;		
Matches 580;	Conservative 0;	Mismatches 120;	Indels 24;	Gaps 4;
Qy	335	TATGATCATCTCTCTTTT		
Db	3310	TTTATATAAGACTTTT		
Qy	395	TCGCCACAGGCTGAGTGCAGTGGCGGGATCTCGGCTCAGTGCAGAGCTTCCGCCCTCCCGGGT		
Db	3250	TCGCCACAGGCTGAGTGCAGTGGCGGGAATCTCGGCTCAGTGCAGAGCTTCCGCCCTCCCGGGT		
Qy	455	TCACGCCATTCTCCTGCTCAGGCTCCCAAGTAGCTGGGACTACAGGCGCCCGCCACTAC		
Db	3190	TCACGCCATTCTCCTGCTCAGGCTCCCAAGTAGCTGGGACTACAGGCGCCCGCCACTAC		
Qy	515	GCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGTTTACCGTTTTAGCCGGGATG		
Db	3130	GCCCGGCTAA-TTTTTGTATTTTTAGTAGAGACGGGTTTACCGTTTTAGCCGGGATG		
Qy	575	GCCTCGATCTCTGACTCGTGTATCGCCCGCTCGGCCCTCCCAAAGTGTGGGATTACA		
Db	3071	GTCTCGATCTCTGACTCGTGTATCGCCCGCTCGGCCCTCCCAAAGTGTGGGATTACA		
Qy	635	GGCGTGAGCCACCGCGCCGCGCATGATCATCTTCTTGACTATGCTGATGTGACAAAGTAC		
Db	3011	GGCGTGAGCCACCGCGCCGCGCC-----TTATAAAGACTTTTATAGATAAAACAAGAAT		
Qy	695	CTAAGGCATCAGACTCTACCTTTAAATATGACAGTTTGGGCCACGACCGTGGCTCATG		
Db	2958	CTATATCTCTAAATGTATG,----TATTAAAGATGAACCCCGGCTGGGCGCAGTGGCTCACA		
Qy	755	CCGTGTAATCCAGCACTTTGGGAGGACAGAGTGGGTGAATCATCTTGAGGCCACGAGGTTCG		
Db	2902	CCGTGTAATCCAGCACTTTGGGAGGCGAGGCGGGTAGATCATCTTGAGTTCAGGAGTTCG		
Qy	815	AGACCAAGCTGGCCAAACATGGTGAACCTCTGTCTTTACTAAAAA		
Db	2842	AGACCAAGCTGGCCAAACATGGTGAACCTCTGTCTTTACTAAAAA		

[illegible]

RESULT 7

```

US-09-949-016-15787/c
; Sequence 15787, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15787
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15787

```

Qy	422	ATCTCGGCTCACTGCAGACTCCGGCTCCGGGTTCACGCCAATTCTCCTGCCTCACGCTCC	481
Db	32423	ATCTCGGCTCACTGCAGACTCCGGCTCCGGGTTCACGCCAATTCTCCTGCCTCACGCTCC	32364
Qy	482	CAAGTAGCTGGGACTACAGGCGCCGCCCNACTACGCCCGCGTAA--TTTTTTTGTAATTTTA	540
Db	32363	CAAGTAACAGGGACTACAGGCGCCGCCCNACCACCGCTGGCTAAATTTTTTTTGTAATTTTA	32304
Qy	541	GTAGAGACGGGGTTTCACCGTTTTTAGCCGGGATGCGCTCGATCTCCTGACCTCGTGATCC	600
Db	32303	GTAGAGACGGGGTTTCACCTGTGTGTAGCCAGGAAGTCTTGATCTCCTGACCTCGTGATCC	32244
Qy	601	GCCCGCTCGGCTCCCAAAGTCTGGGATTAACAGCGGTGAGCCACCGCGCCCGCCCATG	660
Db	32243	ACCCACTCGGCTCCCAAAGTCTGGGATTAACGGCGGTGAGCCACCGCGCCCGCCGAGC	32184
Qy	661	ATCATCTTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTACCCTTTA	720
Db	32183	CATGTCGGGATTCGT-TGATGTTCTTCACTTACCTGAAGGCCACCGTGGAAAGAGGCTT--	32127
Qy	721	AATATGCAAGTTTGGGCCAGGCACGTGGCTCATGCTGTAAATTCAGACACTTTTGGGAGGC	780
Db	32126	-----AGACTTGGCTGTGACGGTGGCTCACACTGTCTCTTAGCACCTTTGGGAGGC	32074
Qy	781	AGAGTGGGTGAATCACTTTGAGGCCAGGAGTTTGAGACACAGGCTCGCCAAACATGTGAAA	840
Db	32073	CGAGGCAGGTAGATCACTTGAGGCCAGGAGTTTCGAGACAAGTCTTGACCAACATGTGAAA	32014
Qy	841	CTCTGTCTTTACTAAAAAATAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCGTGGGGCAC	900
Db	32013	CCCCGCTCTACT------AAACTACAAAAAATAGCTGGCGGTGCTGTGTCAC	31966
Qy	901	ACCTGTAATCCCAGCTATCTCGAGGCTTGAGGCACGAGTGCACATTGACCTCGGAGCG	960
Db	31965	ACCTGTAATCCCAGCTACTTCAGAGGCTGAGGCAGAGAATCTCTTGAACTCAGGAGCA	31906
Qy	961	GAGGTTGCAGTGGCGGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGACGAAGACT	1020
Db	31905	GAGGTTGCAGTGAGCCAAGATTGCGCCACTGCACTCCAGCCTGGGTGACAGAGCCAGACT	31846
Qy	1021	CTGCTCAAAATAATAA 1038	
Db	31845	CTGCTCCAAAAACAAA 31828	

RESULT 8
US-09-949-016-15788/c
; Sequence 15788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIORITY APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIORITY APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIORITY APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15788
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15788

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 2001.63 Seconds

(without alignments)

15667.668 Million cell updates/sec

Title: US-09-936-271C-13_COPY_1_5000

Perfect score: 5000

Sequence: 1 gggccagagtgaggcaag.....tgatccacagtggtgctc 5000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched:

6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	9.4	525	16	US-10-029-386-13540
2	440.2	8.8	4154	13	US-10-087-192-1642
3	419	8.4	91352	18	US-10-300-611-4
4	418.4	8.4	115935	19	US-10-775-169-241
5	417.6	8.4	26371	13	US-10-087-192-1450
6	417.6	8.4	91760	13	US-10-087-192-844
7	417.6	8.4	136726	17	US-10-085-117-244

8	416.2	8.3	379652	21	US-10-481-613-71	Sequence 71, Appl
9	415.2	8.3	17397	9	US-09-764-869-1945	Sequence 1945, Ap
10	415.2	8.3	17397	14	US-10-091-504-1945	Sequence 1945, Ap
11	415.2	8.3	17397	17	US-10-227-577-1945	Sequence 1945, Ap
12	415.2	8.3	19334	9	US-09-764-869-1943	Sequence 1943, Ap
13	415.2	8.3	19334	14	US-10-091-504-1943	Sequence 1943, Ap
14	415.2	8.3	19334	17	US-10-227-577-1943	Sequence 1943, Ap
15	415.2	8.3	19345	9	US-09-764-869-1944	Sequence 1944, Ap
16	415.2	8.3	19345	14	US-10-091-504-1944	Sequence 1944, Ap
17	415.2	8.3	19345	17	US-10-227-577-1944	Sequence 1944, Ap
18	415.2	8.3	29001	19	US-10-317-270-11	Sequence 11, Appl
19	414.8	8.3	12970	10	US-09-764-891-7689	Sequence 7689, Ap
20	414.6	8.3	111084	20	US-10-723-860-1627	Sequence 1627, Ap
21	414.2	8.3	212231	13	US-10-087-192-1126	Sequence 1126, Ap
22	412.8	8.3	3294	13	US-10-027-632-115723	Sequence 115723, Ap
23	412.8	8.3	3294	13	US-10-027-632-115724	Sequence 115724, Ap
24	412.8	8.3	3294	17	US-10-027-632-115724	Sequence 115724, Ap
25	411.2	8.2	13170	19	US-10-741-601-5744	Sequence 5744, Ap
26	411.2	8.2	13170	21	US-10-741-601-5744	Sequence 17903, A
27	410.4	8.2	10445	10	US-09-764-891-6380	Sequence 6380, Ap
28	409.6	8.2	13328	21	US-10-741-600-17627	Sequence 17627, A
29	409.6	8.2	13328	21	US-10-741-600-17720	Sequence 17720, A
30	409.6	8.2	17303	21	US-10-741-600-17966	Sequence 17966, A
31	409.6	8.2	44801	21	US-10-364-505-6	Sequence 6, Appli
32	409.6	8.2	50000	17	US-10-681-199-6	Sequence 6, Appli
33	409.6	8.2	50000	19	US-10-719-993-6883	Sequence 6883, Ap
34	407.6	8.2	187844	20	US-10-087-192-484	Sequence 484, App
35	407.6	8.2	196686	13	US-09-764-891-7689	Sequence 6890, Ap
36	406.2	8.1	12970	10	US-09-764-891-7689	Sequence 6890, Ap
37	405.8	8.1	4433	14	US-10-091-572-668	Sequence 13, Appl
38	405.8	8.1	4433	14	US-10-175-492-13	Sequence 4, Appli
39	405.8	8.1	126001	17	US-10-004-113-4	Sequence 4, Appli
40	405	8.1	32767	16	US-10-394-948-4	Sequence 5704, Ap
41	405	8.1	32767	17	US-10-741-601-5704	Sequence 17804, A
42	405	8.1	67088	19	US-10-087-192-916	Sequence 916, App
43	405	8.1	67088	21	US-09-795-668-1	Sequence 1, Appli
44	403.8	8.1	104062	13		
45	403.6	8.1	1503841	9		

ALIGNMENTS

RESULT 1

US-10-029-386-13540
; Sequence 13540, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13540
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.83
; OTHER INFORMATION: SWISSPROT HIT: Q9Y337, EVALUE 2.00e-34
; OTHER INFORMATION: NT HIT: AFI35028.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE388198.1, EVALUE 0.00e+00

US-10-029-386-13540

Query Match 9.4%; Score 472; DB 16; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.5e-102;


```

; SEQ ID NO 1450
;
; LENGTH: 28371
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
US-10-087-192-1450

```

Query Match	8.4%	Score 417.6;	DB 13;	Length 26371;
Best Local Similarity	78.0%;	Pred. No. 6.6e-89;		
Matches 551;	Conservative	0;	Mismatches 129;	Indels 26;
				Gaps 3;

[illegible]

RESULT 6

US-10-087-192-844
; Sequence 844, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22

```

; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 844
; LENGTH: 91760
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(91760)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-844

```

Query Match	8.4%;	Score 417.6;	DB 13;	Length 91760;
Best Local Similarity	76.8%;	Pred. No. le-88;		
Matches 576;	Conservative	0;	Mismatches 139;	Indels 35; Gaps 4;

[illegible]

RESULT, T 7

US-10-085-117-244/c

US-10-085-117-244/c

Sequence 244, Application US/10085117
Publication No. US20030232334A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 52945200121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 244
LENGTH: 136726
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variation
LOCATION: (1)...(136726)
OTHER INFORMATION: n = any nucleotide
US-10-085-117-244

Query Match 8.4%; Score 417.6; DB 17; Length 136726;
Best Local Similarity 80.5%; Pred. No. 1.2e-88;
Matches 546; Conservative 0; Mismatches 109; Indels 23; Gaps 4;

QY 362 TTTTCTTTTGTGAGAGGAGTCTCGCTGTGTCGCGCCAGGCTGGAGTGGTGGCGG 421
DB 40483 TTCTTTTGTGAGAGGAGTCTCGCTGTGTCGCGCCAGGCTGGAGTGGTGGCG 40424

QY 422 ATCTCGGCTCACTCAAGCTCCGCTCCGCGTTCAGCCATCTCTGCTCAGCTCC 481
DB 40423 ATCTCGGCTCACTCAAGCTCCGCTCCGCGTTCAGCCATCTCTGCTCAGCTCC 40364

QY 482 CAAGTAGCTGGAGTACAGGCGCGCCGCTACGCTACGCGCGGTAA-TTTTTTGTATTTTA 540
DB 40363 CAAGTAACAGGACTACAGGCGCGCCGCTACGCTACGCGGTAA-TTTTTTGTATTTTA 40304

QY 541 GTAGAGAGGCGTTTACCGTTTGTAGCGGATGGCTCGATCTCTGATCTGATCC 600
DB 40303 GTAGAGAGGCGTTTACCGTTTGTAGCGGATGGCTCGATCTCTGATCTGATCC 40244

QY 601 GCCCGCTCGGCTCCCAAGTGTGGATACAGGCGGTGAGCCACGCGCGCGCCATG 660
DB 40243 ACCACCTCGGCTCCCAAGTGTGGATACAGGCGGTGAGCCACGCGCGCGCCAGC 40184

QY 661 ATCATCTTCTGACTATGCTGATGACAGTAAAGCTACATCAAGCTCACTCCCTTTA 720
DB 40183 CATCTCGGATCTGT-TGATGTCTTCACTTCACTGAGGCGCAGTGGAAAGGCTT-- 40127

QY 721 AATATGAGTGTGGGCGGACCGTGGCTCATCGCTGTAATTCAGACATTTTGGGAGGC 780
DB 40126 -----AGACTTGGCTGTACCGTGGCTCACACCTGTATCTTAGACATTTGGGAGGC 40074

QY 781 AGAGTGGTGAATCACTTAGGCGGAGGTTGAGACAGCTGGCGCAACATGGTAAA 840
DB 40073 CGAGGAGGTAGATCACTTAGGCGGAGGTTGAGACAGTCTGACCAACATGGTAAA 40014

QY 841 CTCTGTCTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900
DB 40013 CCCCCTCTACT-----AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39966

QY 901 ACCTGTAATCCAGCTATGCTGAGGCTGAGGCGAGAGTCACTTGAACCTCGAGGCG 960
DB 39965 ACCTGTAATCCAGCTATGCTGAGGCTGAGGCGAGAGTCACTTGAACCTCGAGGCG 39906

QY 961 GAGGTGAGTGGCGGAGTACATCAACCGCTCCAGCTGGCGGAGAGCAAGT 1020
DB 39905 GAGGTGAGTGGCGGAGTACATCAACCGCTCCAGCTGGCGGAGAGCAAGT 39846

QY 1021 CTGTCTCAATAAATAA 1038
||||| ||| |||

Db 39845 CTGTCTCCAAACAAAA 39828

RESULT 8
US-10-481-613-71
Sequence 71, Application US/10481613
Publication No. US20050085627A1
GENERAL INFORMATION:
APPLICANT: Zhang, Youming
APPLICANT: Moffatt, Miriam
APPLICANT: Cookson, William
APPLICANT: Tinsley, Jon
TITLE OF INVENTION: Atopy
FILE REFERENCE: 16721-0003US1 / P32688WO/KVC
CURRENT APPLICATION NUMBER: US/10/481,613
CURRENT FILING DATE: 2003-12-19
PRIOR APPLICATION NUMBER: PCT/GB02/02859
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: GB 0115211.5
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: GB 0115212.3
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: GB 0115213.1
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 326
SOFTWARE: PatentIn version 3.1
SEQ ID NO 71
LENGTH: 379652
TYPE: DNA
ORGANISM: Homo sapiens
US-10-481-613-71

Query Match 8.3%; Score 416.2; DB 21; Length 379652;
Best Local Similarity 78.8%; Pred. No. 3.8e-88;
Matches 559; Conservative 0; Mismatches 98; Indels 54; Gaps 3;

QY 337 TGATCATCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 396
DB 312766 TTAATCATGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 312825

QY 397 -GCCCAGGCTGAGTGCAGTGGCGGATCTCGCTCACTCAAGCTCCGCTCCCGGTT 455
DB 312826 AGCCAGACTGGAGTGTAGTGGCACCATCTCGCTCACTCAAGCTCCGCTCCCGGTT 312885

QY 456 CAGCCATCTCTGCTCAGCTCAGCTCCCAAGTGTGGACTACAGGCGCGCGCCATCAG 515
DB 312886 CACACCATCTCTGCTCAGCTCAGCTCCCAAGTGTGGACTACAGGCGCGCGCCACG 312945

QY 516 CCGGCTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 575
DB 312946 CCGGCTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 313005

QY 576 CTTCTGATCTCTGACCTCGTGATCCGCGCTCGGCTCCCAAGTGTGGAGTTACAG 635
DB 313006 TCTGATCTCTGACCTCGTGATCCGCTCGGCTCCCAAGTGTGGAGTTACAG 313065

QY 636 GCGTGAGCCACCGCGCGCGCTGATCATCTTTCTTTGATGCTGATGTGCAAGTACC 695
DB 313066 GCGTGAGCCACCGCGCGCGCTGATCATCTTTCTTTGATGCTGATGTGCAAGTACC 313098

QY 696 TAAAGCATCAGACTTACCCCTTTAATATGAGTTGGGCGAGGACCGTGGCTCATGC 755
DB 313099 -----TCTTAAATGTGTGTATGCCAGGCTCGGAGACTCAC 313137

QY 756 CTGTAATTCAGACATTTGGGAGGAGGTTGGTGAATCACTTGGAGCCAGAGTTTGA 815
DB 313138 ATGTAATTCAGACATTTGGGAGGAGGTTGGTGAATCACTTGGAGCCGAGTTTGA 313197

QY 816 GACCAGCTGGCCAAACATGTGTAACCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 875
DB 313198 GTCCAGTCTGCCCAACATGTGTAACCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 313252

QY 876 AATCAGCGGCTGTCTGTGGGCGACCTGTAATCCAGCTATGCTGAGGCTGAGGCAC 935
||||| ||| |||

THIS PAGE BLANK (uspto)

Result No.	Query No.	Score	Query		DB	ID	Description
			Match	Length			
C	1	397.6	8.0	2971	8	AF101960	AF101960 AF101960
	2	373.8	7.5	2371	8	AF101960	AF101960 AF101960
	3	365	7.3	1715	3	CS952231	CS952231 full-length
	4	362.16	7.3	2429	3	HSW8053366	AL834319 Homo sapi
	5	337.4	6.7	652	1	AU120416	AU120416 Homo sapi
	6	337.2	6.7	1641	3	BC033224	BC033224 Homo sapi
	7	336	6.7	2772	3	BC038630	BC038630 Homo sapi
	8	334.4	6.7	698	5	EM999040	EM999040 UI-H-D10-
	9	331.8	6.6	5528	9	AB012089	AB012089 Homo sapi
C	10	329	6.6	381	8	CC061638	CC061638 MUGQ CH25
	11	325.2	6.5	652	1	AU120416	AU120416 Homo sapi
	12	324	6.5	701	2	BE744242	BE744242 601576428
	13	324	6.5	5511	3	HSW807323	EX647179 Homo sapi
	14	322.8	6.5	6934	3	BC38830	EX640943 Homo sapi
	15	321.6	6.4	2772	3	HSW807092	BC038630 Homo sapi
	16	318.6	6.4	5528	9	AB012089	AB012089 Homo sapi
	17	315.6	6.3	709	9	AG010328	AG010328 Homo sapi
	18	315	6.3	710	6	CA428305	CA428305 UI-H-DFO-
C	19	312.6	6.3	542	1	AI679782	AI679782 tu76f10. x
	20	312.2	6.2	495	1	AI963720	AI963720 wr65b04. x
	21	311.8	6.2	2429	3	HSW805366	AL834319 Homo sapi
	22	311.4	6.2	3491	3	BC039100	BC039100 Homo sapi
	23	309.6	6.2	1910	3	BC035771	BC035771 Homo sapi
	24	309.4	6.2	3491	3	BC039100	BC039100 Homo sapi

ACCESSION CR592231
VERSION CR592231.1 GI:50473038
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1715)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source . Location/Qualifiers
1..1715
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CS0D1023YE11"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match	7.3%;	Score 365;	DB 3;	Length 1715;
Best Local Similarity	71.5%;	Pred. No. 5.1e-23;		
Matches	551;	Conservative	0;	Mismatches 135; Indels 85; Gaps 2;

```

QY      TTTTTTTTTTTTTTTTTTTTTTTTGAGCGGAGTCTCGCTCTGTCGCCCCAGGCTGGAGTGCAG 414
Db      TTTTTTTTTTTTTTTTTTTTTTTTGAGTCAAGTCTTGCTCTGCGCCCCAGGCTGCAGTGCAG 760

QY      TGGCGGGATCTCGGCTCACGTCCGCTCCCGGCTCCCGGGTTTCAGCCATTCTCTCTGCGCTC 474
Db      TGGCCCGATCTCGGCTCACTGTGAGCTCTGCTCTCGGTTTCATGCCATTCTCTCTGCGCTC 820

QY      AGCTCCCAAGTAGCTGGACTACAGCGGCCCGCCA CTACGCCGGCTAAATTTTGTGTA 534
Db      AGCTCCCAAGTAGCTGGACTACAGCGACCACCA CTACGCCCGCCGCTAAATTTTGTGTA 880

QY      TTTTATAGTAGAGCGGGGTTTCACCGTTTATAGCGGATGGCTCGATCTCTGACCTCG 594
Db      TTTTATAGTAGATGGGGTTTTCACCGTTTAGCCATGATGGTCTCCATCTCTGACCTTG 940

QY      TGATCCCGCCCGCTTCGGCTCCCAGTGTCTGGGATTACAGGGGTGAGCCACCGCGCCG 654
Db      TGACCTGCTTCGCTCGGCCCGCCCAAAGGGCTGGGATTACAGGGGTGAGCCACCGCACCA 1000

QY      GGCA----- 658
|||||
1001   GCCACTTTCTTGATAAAGAAGCACTGACTATGGGCTGTTCTGSCCAGTTTATAGAG 1060

QY      -----TGATCATCTTCTTGACTATGCTGATGTGACAAGTACCTAAGCC 702
Db      GCTAAGCACCCCTTGTTCCCTTCTGTCTGCTGAACAGGCATTTTGACTTATAGGGTCTAATG 1120

QY      ATCAGACTCTACCCCTTTAAATATUCAGCTTTGGGCGAGGCACCGTGGCTCATGCTGTAAT 762
Db      TTATGCAATTAAAGTTTAAGTCTCCACTTAGGCCTGGCGTGGTGGCTCAGGCTGTAAT 1180

QY      TCCAGCACTTTGGGAGGCAGAGGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGC 822

```



```
RESULT 6
BC033224/c
LOCUS Homo sapiens, clone IMAGE:5022014, mRNA. linear HTC 27-JUN-2002
DEFINITION BC033224
ACCESSION BC033224
VERSION BC033224.1 GI:21619936
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1641)
Direct Submission
Submitted (25-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranda Teai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 43 Row: J Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
Location/Qualifiers
1..1641
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5022014"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
ORIGIN
Query Match 6.7%; Score 337.2; DB 3; Length 1641;
Best Local Similarity 75.3%; Pred. No. 1.1e-19;
Matches 500; Conservative 0; Mismatches 128; Indels 36; Gaps 5;
QY 348 TTTTTCAGTGGGGGATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTTCAGGCCATTCTC 467
Db 1118 TTTTTCAGTGGGGGATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTTCAGGCCATTCTC 1059
QY 408 AGTCAGTGGGGGATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTTCAGGCCATTCTC 467
Db 1058 AGTCAGTGGGGGATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTTCAGGCCATTCTC 999
QY 468 CTGCTCAGCTCCCAAGTAGCTGGGACTACAGGGCCCGGCACATACGCCCGGCTTAATT 527
Db 998 CTGCTCAGCTCCCAAGTAGCTGGGACTACAGGGCCCGGCACACACACACCGGACTAA-TT 940
QY 528 TTTTTCATTTTAGTAGAGCGGGGTTTCACCGTTTTCAGCGGGATGGCTCGATCTCT 587
Db 1118 TTTTTCATTTTAGTAGAGCGGGGTTTCACCGTTTTCAGCGGGATGGCTCGATCTCT 587

Db 939 TTTTTCATTTTCAGTAGAGACAGGGTTTCACCTTGTAGCCAGGATGCTCGATCTCT 880
QY 588 GACCTCGTGTATCCCGCCCTCGGCTCCCAAAAGTGTGGGATTACAGCGGTGAGCCACC 647
Db 879 GACCTCATGATCCACCTCGCTCGGCTCCCAAAAGTGTGGGATTACAGGCATGAGCCACC 820
QY 648 GCGCCCGGCATGATCATCTTCTTGTGATCTATCTGATGTGACAGTACCTAAAGCCATCAG 707
Db 819 ATGCCAGCCCTAAACTCTCAATCTAAATTTGGACAATTTTAAAGATATTTTACTCCAGA- 761
QY 708 ACTCTACCTTTTAAATATGCAATTTGGCCAGGACCGCTCATGCTGTGTAATTCAG 767
Db 760 -----CTTATAAACCAAGCTACAGCCAGGACAGTGTGTCAGCATGTAATCCAG 709
QY 768 CACTTTGGGAGGACAGGTGGTGAATCATTGAGGCCAGGAGTTTGAGACAGCCTGGC 827
Db 708 CACTTTGGGAGGCGGAGGT-GGTGATCACTTGAGGCCAGGAGTTTCAAGACAGCAGCCTGGG 650
QY 828 CAACATGCTGAACCTCTCTTTACTTAAATAAAAAAAAAAAAAAAAAAAAAATCAGCCGG 887
Db 649 CAACAGAGCGAGACTCAAACTCTACAGAAATTTTAAATAAATTA-----GCCGG 602
QY 888 TGTCTGTGGGCACACCTGTATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTG 947
Db 601 ACGTGGGACCCCGCTGTAGTTCCAGCTACTCAGAGG-----ATTGCTTG 555
QY 948 AACCTGTGAGCGGAGGTTGTCAGTGGGCCGAGATCATACACCGCCCTCCAGCCTGGGCG 1007
Db 554 AACCTGGAGTTCAAGGTTGTCAGTGGCGTGAGCATCCCACTCCAGCCTGGGCTG 495
QY 1008 ACAG 1011
Db 494 ACAG 491

RESULT 7
BC038630/c
LOCUS Homo sapiens, Similar to hypothetical protein FLJ20489, clone 2772 bp mRNA linear HTC 04-MAR-2003
DEFINITION IMAGE:5263792, mRNA.
ACCESSION BC038630
VERSION BC038630.1 GI:24116283
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2772)
Direct Submission
Submitted (15-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 73 Row: i Column: 4
This clone has the following problem: retained intron.
Location/Qualifiers
1..2772
FEATURES
source
```

	BM999040.1	GI:19723941	EST.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 698) NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapsb-r@mail.nih.gov Tissue Procurement: Dr. Jose Mercuende cDNA Library preparation: Dr. M. Bento Soares, University of Iowa DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov The following repetitive elements were found in this cDNA sequence: 11-315, >ALU (matched complement) 386-662, >ALU Seq primer: M13 FORWARD POLYA=Yes			
VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
FEATURES	source	Location/Qualifiers						
		1..698	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="IMAGE:5881836"	/tissue_type="Lung Focal Fibrosis"	/dev_stage="Adult"
			/lab_host="DH10B	(life Technologies)"	/clone_lib="NCI CGAP D10"	/notes="Organ: Lung; Vector: pTV73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP D10 is a cDNA library containing the following tissue(s): A pool of lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTV73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ATACGGCCTC.		
			TAG_TISSUE=lung with fibrosis	TAG_LIF=UI-H-D10	TAG_SEQ=ATACGGCCTC			
ORIGIN	Query Match	6.7%; Score 334.4; DB 5; Length 698;	Best Local Similarity	71.1%; Pred. No. 3.1e-19;	Mismatches 517; Conservative	0; Mismatches 164; Indels	Gaps	4
Qy	358	TTTTTTTTTTTTTTTTTTTCGAGACGGAGTCTCGCTCTGTGCCCGAGCTCGAGTGCA GTGG 417	Db	1	TTTTTTTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTCAACCAGTCTCGAGTGCA GTGG 60	Qy	418	CGGGATCTCGGCTCACTGC AAAGCTCCGCTCCCGGGTTTCACGCCAATCTCTCCCTCACG 477
Db						Db	61	CGCGATCTCGGCTCACTGC AGGCTCTGCCTCTCTGGGTTTCAGCCATTCTCTGTCTCACG 120
Qy						Qy	478	CTCCCAGTAGCTGGGACTACAGGCCCGCCGCACATAAGCCCGGCTTAATTT----- 526
Db						Db	121	CTCCCCAGCAGCTGGGACAACAGGGCCACAGCCCGCCGCTAAATTTTGTGTGTGT 180
Qy						Qy	527	-----TTTTTGTATTTTGTAGACGGGGTTTACCGTTTTAGCCGGATGGCCTCGAT 582
Db						Db	181	GTGTGTGTGTATTTTGTAGTAGACGGGGTTTACCGTTGTGGCCAGGATGTCTCAAC 240
Qy						Qy	583	CTCCTGACCTCGTGATCCGCCCGCTCGGCCTCCCAAAGTGTGGGAATACAGGCGTGAG 641

RESULT 8	
BM999040	
LOCUS	698 bp mRNA linear EST 17-JUN-2002
DEFINITION	UI-H-D10-avn-j-13-0-UI.s1 NCI CGAP_D10 Homo sapiens CDNA clone
	IMAGE:5881836 3', mRNA sequence.
ACCESSION	BM999040

```
Db 241 CTCCTGACCTCGTGATCTGCCCGCTCCACCTCCCAAGTGTGGATTACAGGGTGAG 300
Qy 643 CCACCGCGCGCCGATGATCATCTTCTTGACTATGCTGATGTGCAAGTACCTAAAGCC 702
Db 301 CCACCATGCTCGCCCATTTATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 360
Qy 703 ATCAGACTCTACCCCTTTAAATATGAGTTTGGGCCAGGCACCGTGGCTCATGCTGTAAT 762
Db 361 GAAGACTGACGAGACTACT-----GGTTGGCGGGCGTGGTGACTCATATCTGAAT 414
Qy 763 TCAGCACTTTTGGGAGCGAGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGC 822
Db 415 CCAGTACTTCAGGAAGCAAGAGGGTGAATCTTTTGGCCCGAGAAATTCGAGACCAGG 474
Qy 823 CTGGCCACATGGTGAATCTGTCTTTACTTAAATAAAAAAATAAAAAAATAAATCAG 882
Db 475 CTGGCCACATGGCAAAACCCCTGTCTTACAAAAAATA-----513
Qy 883 CGGGTGTCTGGGCGACACCTGTAATCCAGCTATGCTGGAGGCTGAGGCAGCAGAGTC 942
Db 514 ---GGTGTGTGCCACAGCCCTGTGGTCCAGCTACTCGGGAGGCTGAGGTAATAGGCTC 570
Qy 943 ACTTGAACCTTGGAGCGAGGTTGAGTGGCGGAGATCAATCACCGCCCTCCAGCT 1002
Db 571 ACTAGAACCTGGAGGTTAAGCTCGAGTGGAGCGAGATCTGCGCACTGTACTCCAGCT 630
Qy 1003 GGGCGACAGAGCA-AGACTCTGTCTCAATAAATAAATAAACAACGACAGCAGTTTG 1061
Db 631 GAGTGGCAGAGCANACCTCTCTCANAAATAATTTTANNAAGAAAAAATAAAAAAAC 690
Qy 1062 TTGTACC 1068
Db 691 TCGTGCC 697
```

RESULT 9

AB012089 Homo sapiens gene, genomic survey sequence. 528 bp DNA linear GSS 21-MAY-1998

LOCUS AB012089

DEFINITION Homo sapiens gene, genomic survey sequence.

ACCESSION AB012089

VERSION AB012089.1 GI:3149956

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

Kurashashi, H., Sakamoto, M., Ono, J., Honda, A., Okada, S. and

Nakamura, X.

Molecular cloning of the chromosomal breakpoint in the Lif1 gene of

a patient with isolated lissencephaly and balanced t(8:17)

Hum. Genet. (1998) In press

2 (bases 1 to 528)

Kurashashi, H.

Direct Submission

Submitted (09-MAR-1998) Hiroki Kurashashi, Biomedical Research

Center, Osaka University Medical School, Division of Clinical

Genetics, Department of Medical Genetics; 2-2 Yamadaoka, Suita,

Osaka 565-0871, Japan (E-mail: kurashashi@genetics.med.osaka-u.ac.jp,

Tel: 81-6-879-3381, Fax: 81-6-879-3389)

Location/Qualifiers

1..528

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17p13.3"

ORIGIN

Query Match 6.6%; Score 331.8; DB 9; Length 5528;

Best Local Similarity 62.0%; Fred. No. 1.3e-19;

Matches 629; Conservative 0; Mismatches 347; Indels 39; Gaps 5;

LOCUS CC061638

381 bp DNA linear GSS 28-APR-2003

RESULT 10

CC061638/c

LOCUS

DEFINITION MUCQ_CH252P004Q3T7.M05_CD318_034_CHORI-252 Vervet Monkey Library
Cercopithecus aethiops genomic clone CH252-4B9, genomic survey
sequence.

ACCESSION CC061638
VERSION CC061638.1 GI:29790901
KEYWORDS GSS.
SOURCE Cercopithecus aethiops (African green monkey)
ORGANISM Cercopithecus aethiops
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Cercopithecus.
1 (bases 1 to 381)
REFERENCE Minhas, R., Zhang, X., Dore, C., Villeneuve, A., Lepage, P.,
Forgetta, V., McKee, K., Ophoff, R.A., Fairbanks, L.A., Freimer, N.B.,
Ervin, F.R., Palmour, R.M., Hudson, T.J. and Dewar, K.
UCLA/MUGO/St-Kitts Vervet Monkey Mapping Project
Unpublished (2003)
CONTACT: Dewar, K.
McGill University and Genome Quebec Innovation Centre
McGill University
740 Dr. Penfield Room 7214, Montreal, QC, Canada, H3H 1A4
Tel: 514 398 3311 x00089
Fax: 514 398 1795
Email: ken.dewar@mcgill.ca
Plate: 4 row: B column: 9
Seq primer: T7 : TAATACGACTCACTATAGG
Class: BAC ends.

FEATURES
Location/Qualifiers
1..381
/organism="Cercopithecus aethiops"
/mol_type="genomic DNA"
/db_xref="taxon:9534"
/clone="CH252-4B9"
/sex="male"
/cell_type="White blood cell"
/dev_stage="Adult"
/clone_lib="CHORI-252 Vervet Monkey Library"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
Constructed by Michael Nefedov in Pieter de Jong's
laboratory at BACPAC Resources, Children's Hospital in
Oakland Research Institute."

ORIGIN
Query Match 6.6%; Score 329; DB 8; Length 381;
Best Local Similarity 93.2%; Pred. No. 1.3e-18;
Matches 355; Conservative 0; Mismatches 25; Indels 1; Gaps 1;
Qy 1649 GGAGACCCCTCCCATCTCCGGGCCAGGAGGTAGGAGTGACATTCGGACTGGGTGG 1708
Db 381 GGAGTCCCTCCCATCTCTGAGCCAGGAGGTAGGAGTGGTATTCCGACGGGATGG 322
Qy 1709 GGGGTGCTCTGGGGTGGAGATAGGGGAGCAGGAGGAGCTATTGCTAAGCCCGATAGG 1768
Db 321 GGGGTGCTCTGGGGTGGAGATAGGGGAGCAGGAGGAGCGATTGCTAAGCCCAATAGG 262
Qy 1769 CACCTCATTGCCCGGAATGTGCCCGAGGAGCAGTGGGTGTTATTAACCTCAGGCCGGT 1828
Db 261 CACCTCATGGCCCGGAATGTGCCCGAGGAGCAGTGGGTGTTATTAACCTCAGGCCGGT 202
Qy 1829 GCCCAGAGCCAGGAGGAGGAGTGGCCAGGAAGGCACAGGCCCTGAGAAGTCTGCGGCTG 1888
Db 201 GCCCAGAGCCAGGAGGAGGAGTGGCCAGGAAGGTGCGAGGCCCTGAGAAATCCCGGCTG 142
Qy 1889 AGCTGGAGCAAAATCCCCACCCCTACTCTGGGGACAGGGCAAGTGAAGCTGGTGAGG 1948
Db 141 AACTGGAAGCAAAATCCCCACCCCTACTCTGGGGACAGGGCAGGTGAGACTGGGGAGG 82
Qy 1949 GTGGCTCAGCAGGAGGAGGAGGTCTGTGCTGCTGCAACCCACATCTTCTC-T 2007
Db 81 GTGGCTCAGCAGGAGGAGGAGGTCTGTGCTGCTGCAACCCACATCTTCTCTT 22
Qy 2008 GTCCCTCTTGGCCCTGTCTG 2028
|| ||||| ||||| |||||

Db 21 GTTCCCTCTCTCGCCCGCTCTG 1

RESULT 11
AUI20416
LOCUS AUI20416
DEFINITION AUI20416 HEMBB1 Homo sapiens cDNA clone HEMBB100637 5', mRNA
sequence.
ACCESSION AUI20416
VERSION AUI20416.1 GI:10935651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 652)
REFERENCE Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Negai, T., Sugano, S. and
Isoqai, T.
HRI human cDNA project
Unpublished (2000)
CONTACT: Takao Isoqai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
Location/Qualifiers
1..652
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBB100637"
/tissue_type="whole embryo, mainly body"
/dev_stage="embryo, 10 weeks"
/clone_lib="HEMBB1"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 6.5%; Score 325.2; DB 1; Length 652;
Best Local Similarity 74.7%; Pred. No. 1.9e-18;
Matches 489; Conservative 0; Mismatches 148; Indels 18; Gaps 6;
Qy 352 TTTTITTTTTTTTTTTTTTTTTTTTGTGAGACGAGTCTCGCTC-TGTGCGCCAGGTGGAGT 410
Db 3 TCTTTCAACATTTTTTTTTTTTTTTTGTGACAGAGTTTTGCTCTTCCGCTCCAGGCTGGAGT 62
Qy 411 GCAGTGGCGGATCTCGGCTCATTGGAAGTCCCGCTCCCGGTTTCAGCCATTCTCTCTG 470
Db 63 GCATGGCGGATCTCGGCTCGCTGCAACCTCCACCTCTCGGTTTCAAGCAATTTCTCTA 122
Qy 471 CCTCAGCTCCCAAGTAGCTGGGACTACAGCGCCCGCCACTACGCCGGCTAATTTTTT 530
Db 123 CCTCAGCTCCCAAGTAGCTGGGATTAAGCACTGACATCAGCCCGAGCTAA--TTTT 180
Qy 531 TGTATTTTATAGTAGAGCGGGTTTTTACCGTTTTTAGCCGGATGGCTTCGATCTCTGAC 590
Db 181 TGTATTTTATAGTAGAGTGGGTTTTTACCATTGTTGGCAAGGCTGGTCTCAAACTCTGAC 240
Qy 591 CTC--GTGATCCGCGCTCGGCTCCCAAGTGTGGGATTACAGCGTGAAGCCACCG 648
Db 241 CTCAGGTGATCCAACTGCCTCGGCTTCCGAAGTGTGGGATTACAGCGCTGAGCCACCG 300
Qy 649 CGCCCGGCCATGATCATCTTCTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGA 708
Db 301 TGCCTGCGCTTTTATTTTGTGTGTGTGTTGATTCTGACTTGGCTTTTAAATAATC 360
Qy 709 CTCTACCCCTTTAAATATGACAGTTTGGGCCAGGCCCGTGGCTCATGCTGTAATTCAGC 768
|| ||||| ||||| |||||

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 18:09:16 ; Search time 2885.07 Seconds

(without alignments)
16811.954 Million cell updates/sec

Title: US-09-936-271C-13_COPY_6000_7000

Perfect score: 1001

Sequence: 1 gacaaccaagagagcccccag.....gttactgtattggatggcac 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	6	CQ788219 Sequence
2	1001	100.0	11570	6	CQ874885 Sequence
3	1001	100.0	11570	6	CQ874961 Sequence
4	1001	100.0	11570	9	AF135028 Homo sapi
5	1001	100.0	107487	9	AC011483 Homo sapi
6	1001	100.0	230000	9	AF243527 Homo sapi
7	984.2	98.3	217346	2	AC027602 Homo sapi
8	958.8	95.8	200792	2	AC130782 Pan trogl
9	284.6	28.4	142003	9	AC098800 Homo sapi
10	276.8	27.7	177787	9	AC133961 Homo sapi
11	274.4	27.4	138411	9	HSJ71H19 Human DNA
12	262.4	26.2	92882	9	AL355815 Human DNA
13	261.6	26.1	85304	9	HS227L5 Human DNA
14	260	26.0	161014	9	AC087591 Human DNA
15	260	26.0	177876	2	AC025163 Homo sapi
16	256.4	25.6	177562	9	HS269M15 Human DNA
17	253.6	25.3	136649	9	HSJ7635E8 Human DNA
18	253.2	25.3	173840	9	AC092445 Homo sapi
19	253.2	25.3	206819	2	AC016029 Homo sapi

20	252.6	25.2	52216	6	BD012166	BD012166	Isolation
21	252.6	25.2	52216	6	BD128690	BD128690	Method fo
c 22	252.6	25.2	122302	9	AC003982	AC003982	Homo sapi
23	252.6	25.2	220384	2	AC078926	AC078926	Homo sapi
c 24	251.8	25.2	70356	9	AL353794	AL353794	Human DNA
25	251.8	25.2	154664	9	AC025362	AC025362	Homo sapi
26	251.8	25.2	160426	2	AC087697	AC087697	Homo sapi
27	251.8	25.2	162354	2	AC073299	AC073299	Homo sapi
c 28	251.8	25.2	165330	2	AC016011	AC016011	Homo sapi
29	251.8	25.2	179219	9	AC016489	AC016489	Homo sapi
30	250.8	25.1	163453	9	AC099561	AC099561	Homo sapi
c 31	250	25.0	195478	9	AC013409	AC013409	Homo sapi
c 32	249.4	24.9	33768	9	AC078783	AC078783	Homo sapi
c 33	249	24.9	74705	9	AC116339	AC116339	Homo sapi
34	249	24.9	161133	2	AC019295	AC019295	Homo sapi
35	249	24.9	163690	2	AC083781	AC083781	Homo sapi
c 36	249	24.9	169544	9	AC008573	AC008573	Homo sapi
37	249	24.9	179665	2	AC010246	AC010246	Homo sapi
38	249	24.9	195476	9	AC010259	AC010259	Homo sapi
c 39	249	24.9	197810	9	AC090071	AC090071	Homo sapi
c 40	249	24.9	253038	2	AC008930	AC008930	Homo sapi
41	248.8	24.9	193772	9	AC007073	AC007073	Homo sapi
c 42	248	24.8	66987	2	AP002740	AP002740	Homo sapi
c 43	247.2	24.7	153605	2	AC019360	AC019360	Homo sapi
c 44	247.2	24.7	187546	2	AC129236	AC129236	Homo sapi
c 45	246.8	24.7	180939	2	AL596206	AL596206	Homo sapi

ALIGNMENTS

RESULT 1	CQ788219	Sequence 2 from Patent WO2004021008.	11570 bp	DNA	linear	PAT 24-MAR-2004
LOCUS	CQ788219	Sequence 2 from Patent WO2004021008.	11570 bp	DNA	linear	PAT 24-MAR-2004
DEFINITION	CQ788219	Sequence 2 from Patent WO2004021008.	11570 bp	DNA	linear	PAT 24-MAR-2004
ACCESSION	CQ788219	Sequence 2 from Patent WO2004021008.	11570 bp	DNA	linear	PAT 24-MAR-2004
VERSION	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
KEYWORDS	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
SOURCE	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
ORGANISM	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
REFERENCE	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
AUTHORS	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
TITLE	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
JOURNAL	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
FEATURES	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Location/Qualifiers	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
source	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
organism="Homo sapiens"	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
/mol_type="unassigned DNA"	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
/db_xref="taxon:9606"	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004

ORIGIN	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Query Match	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Best Local Similarity	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Matches 1001; Conservative	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
0; Mismatches	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
0; Indels	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
0; Gaps	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Qy	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
1	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
6000	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Qy	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
61	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
6060	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Qy	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
121	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
6120	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
Qy	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
181	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004
6180	CQ788219.1	GI:45723068	11570 bp	DNA	linear	PAT 24-MAR-2004

		/mol_type="unassigned DNA" /db_xref="taxon:9606"			
		ORIGIN			
		Query Match 100.0%; Score 1001; DB 6; Length 11570; Best Local Similarity 100.0%; Pred. No. 9,7e-228; Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	241	1	GACAAACAGAGCCCCCAAGTGAGTCCAGGTTCTTTGATACCGACCCATCTCTGC	60	
Db	6240	6000	GACAAACAGAGCCCCCAAGTGAGTCCAGGTTCTTTGATACCGACCCATCTCTGC	6059	
Qy	301	61	CGCTTCCATCTTTCTCCACTCTCTCATTTGTTGACAGTGCCTTCCCTAAGG	120	
Db	6300	6060	CGCTTCCATCTTTCTCCACTCTCTCATTTGTTGACAGTGCCTTCCCTAAGG	6119	
Qy	361	121	TCCTCCAGTGTGAATATCAGCGTCTTAAGTCAGAAAAGTGCAGAGTCTTACCCGA	180	
Db	6360	6120	TCCTCCAGTGTGAATATCAGCGTCTTAAGTCAGAAAAGTGCAGAGTCTTACCCGA	6179	
Qy	421	181	GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTTCCAGG	240	
Db	6480	6180	GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTTCCAGG	6239	
Qy	241	241	TGAGGACACCTCTCTTTATTTCAGCAGATACACACTGAGTGCCTGCTCGTAACTGGAGC	300	
Db	6240	6240	TGAGGACACCTCTCTTTATTTCAGCAGATACACACTGAGTGCCTGCTCGTAACTGGAGC	6299	
Qy	301	301	GTTCGCAAAATCTGAGAAATCCAGCAATTCGCAAGCAGTCAAGACCCCTGTTCTCAGAG	360	
Db	6300	6300	GTTCGCAAAATCTGAGAAATCCAGCAATTCGCAAGCAGTCAAGACCCCTGTTCTCAGAG	6359	
Qy	361	361	GCTCATACCTTAGAGTGTGTTTGTAGTAAATTAATGCTGAGTCTGTTATGATTC	420	
Db	6360	6360	GCTCATACCTTAGAGTGTGTTTGTAGTAAATTAATGCTGAGTCTGTTATGATTC	6419	
Qy	421	421	CAGTTTTTTAGTACCCACATTAATAACAGGTAAAAAGGCTGGGCGAGTGGCTCACACT	480	
Db	6420	6420	CAGTTTTTTAGTACCCACATTAATAACAGGTAAAAAGGCTGGGCGAGTGGCTCACACT	6479	
Qy	481	481	GTAATCCAGACATTTGGGAGGCTGAGCGAGGAGATCACTTTGGTCAGAGTTTGAGA	540	
Db	6480	6480	GTAATCCAGACATTTGGGAGGCTGAGCGAGGAGATCACTTTGGTCAGAGTTTGAGA	6539	
Qy	541	541	CTAGCCTGGCCAAACATCGCGAACTCTGCTCTCTAAAAAATAACAAAAATAGCTCGC	600	
Db	6540	6540	CTAGCCTGGCCAAACATCGCGAACTCTGCTCTCTAAAAAATAACAAAAATAGCTCGC	6599	
Qy	601	601	ATGTTGGCGGCGCTCTGTAATCTCAGCTGCTCAGGAGCGGAGACACAAGAAATCACTTAA	660	
Db	6600	6600	ATGTTGGCGGCGCTCTGTAATCTCAGCTGCTCAGGAGCGGAGACACAAGAAATCACTTAA	6659	
Qy	661	661	ACCAGAGGTGGAGTTGAGTGCAGTGCATCTGCTCACTCCTCACTCCTCACTGGGAGAC	720	
Db	6660	6660	ACCAGAGGTGGAGTTGAGTGCAGTGCATCTGCTCACTCCTCACTCCTCACTGGGAGAC	6719	
Qy	721	721	AGAGTGACATTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG	780	
Db	6720	6720	AGAGTGACATTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG	6779	
Qy	781	781	TTAACTTTTAATACCCCAATGTATCCCAATACATCAATTTCAAAGTGTAAATATAA	840	
Db	6780	6780	TTAACTTTTAATACCCCAATGTATCCCAATACATCAATTTCAAAGTGTAAATATAA	6839	
Qy	841	841	ACAAATTAATGAGATACCTTTTACATTTCTTTTTCATATTAAGTCTTTGAAAGT	900	
Db	6840	6840	ACAAATTAATGAGATACCTTTTACATTTCTTTTTCATATTAAGTCTTTGAAAGT	6899	
Qy	901	901	GAGTATATGTTATGCTGACAGCACAATCTCAATTTGGAGTCTACATTTTCAGTGTCTC	960	
Db	6900	6900	GAGTATATGTTATGCTGACAGCACAATCTCAATTTGGAGTCTACATTTTCAGTGTCTC	6959	
Qy	961	961	AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC	1001	

RESULT 2
LOCUS CQ874885 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 3 from Patent WO2004077060.
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Diamandis, E.P. and Petraki, C.
TITLE Assay for detection of renal cell carcinoma
JOURNAL Patent: WO 2004077060-A 3 10-SEP-2004;
MOUNT Mount Sinai Hospital (CA)
LOCATION/Qualifiers
FEATURES
1..11570
/organism="Homo sapiens"

Db	6960	AGTAGCCACATGGCTAGCAGTACTGTAATGGATGGCCAC	7000
Db	6960	ATGGTGGCGGCGCTGTAATCTCAGCTGCTCAGGAGGCGGAGACACAAGAAATCACTTAA	6655
RESULT 3			
CO874961			
LOCUS	CO874961	11570 bp	DNA linear PAT 27-SEP-2004
DEFINITION	Sequence 5 from Patent WO2004075713.		
ACCESSION	CO874961		
VERSION	CO874961.1	GI:52748060	
KEYWORDS	Homosapiens (human)		
SOURCE	Homosapiens		
ORGANISM	Homosapiens		
REFERENCE			
AUTHORS	Diamandis,E.P.		
TITLE	Multiple marker assay for detection of ovarian cancer		
JOURNAL	Patent: WO 2004075713-A 5 10-SEP-2004;		
JOURNAL	MOUNT SINAI HOSPITAL CORPORATION (CA)		
FEATURES	Location/Qualifiers		
source	1. .11570		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	100.0%;	Score 1001;	DB 6; Length 11570;
Best Local Similarity	100.0%;	Pred. No. 9,7e-28;	
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GACAAACAGAGCCCAAGGTGAGTGTCCAGGTTCTTCTGATACCAACCACTCTGCG	60
Db	6000	GACAAACAGAGCCCAAGGTGAGTGTCCAGGTTCTTCTGATACCAACCACTCTGCG	6059
Qy	61	CGCTTCCATCTTCTCCACTTCTCATTTGTTCTCTGTTTGACAGTGCACCTTCCCTAAGG	120
Db	6060	CGCTTCCATCTTCTCCACTTCTCATTTGTTCTCTGTTTGACAGTGCACCTTCCCTAAGG	6119
Qy	121	TCCTCAGTGTCTTGAATATCAGGCTGCTAAGTCAGAAAGGTCGAGGATGCTTACCCGA	180
Db	6120	TCCTCAGTGTCTTGAATATCAGGCTGCTAAGTCAGAAAGGTCGAGGATGCTTACCCGA	6179
Qy	181	GACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACTCTTGCCAGG	240
Db	6180	GACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACTCTTGCCAGG	6239
Qy	241	TGAGGACACCTCTCTTTATTCAGCAGATACACACTGAGTGCCCACTCGGTAAACATGAGC	300
Db	6240	TGAGGACACCTCTCTTTATTCAGCAGATACACACTGAGTGCCCACTCGGTAAACATGAGC	6299
Qy	301	GTTCGCCAAATTCAGAAATCCAGCAATTCGCAAGACAGTCAGGACCCCTGTTCTCACAGA	360
Db	6300	GTTCGCCAAATTCAGAAATCCAGCAATTCGCAAGACAGTCAGGACCCCTGTTCTCACAGA	6359
Qy	361	GCTCATAACCTCAGAGTAGTGGTGTGTTAGTAGAAATAATGCTGAGCTGTATGTCATTTC	420
Db	6360	GCTCATAACCTCAGAGTAGTGGTGTGTTAGTAGAAATAATGCTGAGCTGTATGTCATTTC	6419
Qy	421	CAGTTTTTTAGTAGCCACATTAACACAGGTAAACAAAGGCTGGGCGCAGTGGCTCACACCT	480
Db	6420	CAGTTTTTTAGTAGCCACATTAACACAGGTAAACAAAGGCTGGGCGCAGTGGCTCACACCT	6479
Qy	481	GTAATCCAGCAGCTTTGGGAGGCTGAGGCGAGCAGATCACCTTTGCTCAGGAGTTTGAGA	540
Db	6480	GTAATCCAGCAGCTTTGGGAGGCTGAGGCGAGCAGATCACCTTTGCTCAGGAGTTTGAGA	6539
Qy	541	CTAGCTGGCCAACTGCGGAACTCTGCTCTCTAAACAAAAATACAAAAATTTAGCTGGC	600
Db	6540	CTAGCTGGCCAACTGCGGAACTCTGCTCTCTAAACAAAAATACAAAAATTTAGCTGGC	6599
Qy	601	ATGGTGGCGGCGCTGTAATCTCAGCTGCTCAGGAGGCGGAGACACAAGAAATCACTTAA	660

/db_xref="GI:4589283"
/translation="MATARPPMMWVLCALITALLGVTEHVLANNVDSCDPSNTVPS
GSDNLGAGAGEDARSDDSSRIINGSDCMHTOPWOALLRLPNOLYCGAVLVHPOM
LLTAHCKKVRVRLGHYSLSPVYESGQMFQVKSI.PHPGYSHPGSHNDLMILKLN
RRIRPTDVRPINVSSCHPSAGTKLVSGMGWTKSPQVHPFKVLQCLNLSIVLSQKRC
DAYPRQIDDDTFMFCAGDKAGRDSQCGSDGPPVNCVNGSLQGLVSMGDFPCARNRPGVYT
NLCKFTKWIQTIQANS"

ORIGIN

Query Match 100.0%; Score 1001; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 9.7e-228;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAAACAGAGCCCCCAGGTGAGTGCACAGGTTCTTCTTGATACCGACCCATCTCTGC 60
Db 6000 GACAAACAGAGCCCCCAGGTGAGTGCACAGGTTCTTCTTGATACCGACCCATCTCTGC 6059

Qy 61 CGCTTCATCTTCTCCACTTCTCAATCTCTCTGTTGACAGTGCACTTCCTCAAGG 120
Db 6060 CGCTTCATCTTCTCCACTTCTCAATCTCTCTGTTGACAGTGCACTTCCTCAAGG 6119

Qy 121 TCCTCCAGTGCTTGAATATCAGCGTCTTAAGTTCAGAAAGGTGCGAGATGCTTACCGGA 180
Db 6120 TCCTCCAGTGCTTGAATATCAGCGTCTTAAGTTCAGAAAGGTGCGAGATGCTTACCGGA 6179

Qy 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGAGGTAGAGACTCTCTGCCAGG 240
Db 6180 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGAGGTAGAGACTCTCTGCCAGG 6239

Qy 241 TGAGGACACCTCTCTTTATCAGCAGATACACACTGAGTGCCAACTCGGTAACATGAGC 300
Db 6240 TGAGGACACCTCTCTTTATCAGCAGATACACACTGAGTGCCAACTCGGTAACATGAGC 6299

Qy 301 GTTGCCAAATCTGAGATCCAGCAATTCGCAAGCAGTCAAGCCCCCTGTTCTCACAGA 360
Db 6300 GTTGCCAAATCTGAGATCCAGCAATTCGCAAGCAGTCAAGCCCCCTGTTCTCACAGA 6359

Qy 361 GCTCATACCTTAGAGTAGTGTTGTAGTGAATAATGCTGAGTGCTTATGTCATTTC 420
Db 6360 GCTCATACCTTAGAGTAGTGTTGTAGTGAATAATGCTGAGTGCTTATGTCATTTC 6419

Qy 421 CAGTTTTTTAGTAGCCACATTAACACAGGTAAAAAGGCTGGCGCAGTGCTCACACCT 480
Db 6420 CAGTTTTTTAGTAGCCACATTAACACAGGTAAAAAGGCTGGCGCAGTGCTCACACCT 6479

Qy 481 GTAATCCAGCATTGTGGAGGCTGAGCAGCAGATCACCTTTGGTCAGAGTTTGAGA 540
Db 6480 GTAATCCAGCATTGTGGAGGCTGAGCAGCAGATCACCTTTGGTCAGAGTTTGAGA 6539

Qy 541 CTAGCTGGCCAAACATGGCGAACTCTCTCTTAAAAAAAATAACAAAAATTAGCCTGGC 600
Db 6540 CTAGCTGGCCAAACATGGCGAACTCTCTCTTAAAAAAAATAACAAAAATTAGCCTGGC 6599

Qy 601 ATGTGGCGGGCGCTGTAATCTCAGCTGCTCAGGAGCCGAGACACAAGAAATCACTTAA 660
Db 6600 ATGTGGCGGGCGCTGTAATCTCAGCTGCTCAGGAGCCGAGACACAAGAAATCACTTAA 6659

Qy 661 ACCGAGAGGTGGAGTTGCGAGTGAGTGTGTCACCTCACTCCAACTGGGAGAC 720
Db 6660 ACCGAGAGGTGGAGTTGCGAGTGAGTGTGTCACCTCACTCCAACTGGGAGAC 6719

Qy 721 AGAGTGACATTTTCTCTCAAAAAGAAAAAAAACAAAGTAAAAAAGAAACAGGTGAAG 780
Db 6720 AGAGTGACATTTTCTCTCAAAAAGAAAAAAAACAAAGTAAAAAAGAAACAGGTGAAG 6779

Qy 781 TTAACCTTTTAACCAATGATATCCAAATACAATCAATTTCAAAGTGAATTAATAATAA 840
Db 6780 TTAACCTTTTAACCAATGATATCCAAATACAATCAATTTCAAAGTGAATTAATAATAA 6839

Qy 841 ACAATTATGAATGAGATCTTTACATCTTTTCTGTTTTCATATTAAGTCTTTGAAAGT 900
Db 6840 ACAATTATGAATGAGATCTTTACATCTTTTCTGTTTTCATATTAAGTCTTTGAAAGT 6899

Qy 901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGGACTAGCTACATTTCAAGTGCTC 960
Db 6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGGACTAGCTACATTTCAAGTGCTC 6959

Qy 961 AGTAGCCACATGTGGCTAGCAGTACTGATTTGATGGATGGCAC 1001
Db 6960 AGTAGCCACATGTGGCTAGCAGTACTGATTTGATGGATGGCAC 7000

RESULT 5
AC011483/c
LOCUS
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE Unpublished
JOURNAL 2 (bases 1 to 107487)
REFERENCE Direct Submission
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
1 .107487
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clones="CTB-147C22"
64998..65494
/note="NOTE: Shatter libraries failed to resolve
dinucleotide repeat. Unsure number of repeat copies
64998-65494. Forced join 65015."

misc_feature
64998..65494
/note="NOTE: Shatter libraries failed to resolve
dinucleotide repeat. Unsure number of repeat copies
64998-65494. Forced join 65015."

ORIGIN
Query Match 100.0%; Score 1001; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 1.2e-227;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAAACAGAGCCCCCAGGTGAGTGCACAGGTTCTTCTTGATACCGACCCATCTCTGC 60
Db 54220 GACAAACAGAGCCCCCAGGTGAGTGCACAGGTTCTTCTTGATACCGACCCATCTCTGC 54161

Qy 61 CGCTTCATCTTCTCCACTTCTCAATCTCTCTGTTGACAGTGCACTTCCTCAAGG 120
Db 54160 CGCTTCATCTTCTCCACTTCTCAATCTCTCTGTTGACAGTGCACTTCCTCAAGG 54101

Qy 121 TCCTCCAGTGCTTGAATATCAGCGTCTTAAGTTCAGAAAGGTGCGAGATGCTTACCCGA 180
Db 54100 TCCTCCAGTGCTTGAATATCAGCGTCTTAAGTTCAGAAAGGTGCGAGATGCTTACCCGA 54041

Qy 181 GACAGATAGATGACACCATGTTCTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTCTGCCAGG 240

```
Db 54040 GACAGATAGATGACACCACTGTTCTCGCGCGGTGACAAAGCAGGTAGAGACTCTCTGCCAGG 53981
Qy 241 TGAGAGACACCTCTCTTTATTGAGCAGATACACACTGAGTGCCAACTCGGTAAACATGAGGAC 300
Db 53980 TGAGAGACACCTCTCTTTATTGAGCAGATACACACTGAGTGCCAACTCGGTAAACATGAGGAC 53921
Qy 301 GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGACAGTTCAGGACCCCTGTTCTCACAGA 360
Db 53920 GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGACAGTTCAGGACCCCTGTTCTCACAGA 53861
Qy 361 GCTCATACCCCTAGAGTAGTGGTGTGTTAGTAGAAATAATGCTGAGCTGCTTATGTCATTTTC 420
Db 53860 GCTCATACCCCTAGAGTAGTGGTGTGTTAGTAGAAATAATGCTGAGCTGCTTATGTCATTTTC 53801
Qy 421 CAGTTTTTTAGTAGCCACATTAAACAGGTAAAAAGGCTGGCGCAGTGCTGCACACCT 480
Db 53800 CAGTTTTTTAGTAGCCACATTAAACAGGTAAAAAGGCTGGCGCAGTGCTGCACACCT 53741
Qy 481 GTAATCCAGCACCTTTGGAGGCTGAGGAGCGCAGATCACCTTTGGTCAGGAGTTTGAGA 540
Db 53740 GTAATCCAGCACCTTTGGAGGCTGAGGAGCGCAGATCACCTTTGGTCAGGAGTTTGAGA 53681
Qy 541 CTAGCTGGCCCAACATGGCGAAACTCTCTCTCTAAAAAATAACAAAAATTAGCCTGGC 600
Db 53680 CTAGCTGGCCCAACATGGCGAAACTCTCTCTCTAAAAAATAACAAAAATTAGCCTGGC 53621
Qy 601 ATGTTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCCGAGACACAAAGAAATCACTTAA 660
Db 53620 ATGTTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCCGAGACACAAAGAAATCACTTAA 53561
Qy 661 ACCAGGAGGTGGAGTTGCGAGTGCAGTGCAGATCGTGCACCTCACTCAACCTGGGAGAC 720
Db 53560 ACCAGGAGGTGGAGTTGCGAGTGCAGTGCAGATCGTGCACCTCACTCAACCTGGGAGAC 53501
Qy 721 AGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAGCTGAAG 780
Db 53500 AGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAGCTGAAG 53441
Qy 781 TTAACCTTTAATACCCCAATGATCCCAATCAATCAATTTCAAAAGTGAATTAATATAA 840
Db 53440 TTAACCTTTAATACCCCAATGATCCCAATCAATCAATTTCAAAAGTGAATTAATATAA 53381
Qy 841 ACAATTATGAATGAGATCACTTACATCTCTTTCTGTTTTCATATTAAGTCTTTGAAAGT 900
Db 53380 ACAATTATGAATGAGATCACTTACATCTCTTTCTGTTTTCATATTAAGTCTTTGAAAGT 53321
Qy 901 GAGTATATGTTATGCTGAGACACATCTCAATTTGGACTAGCTACATTTTCAGTGTCTC 960
Db 53320 GAGTATATGTTATGCTGAGACACATCTCAATTTGGACTAGCTACATTTTCAGTGTCTC 53261
Qy 961 AGTAGCCACATGTGGCTAGCAGTACTGATTTGGATGGCAC 1001
Db 53260 AGTAGCCACATGTGGCTAGCAGTACTGATTTGGATGGCAC 53220
```

RESULT 6

AF243527/c 230000 bp DNA linear PRI 21-NOV-2000
LOCUS Homo sapiens serine protease gene cluster, complete sequence.

AF243527

AF243527

AF243527.1 GI:11244757

Homo sapiens (human)

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 230000)

Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,

Moss, P., Paepel, B. and Wang, K.

Sequencing and expression analysis of the serine protease gene

cluster located in chromosome 19q13 region

Gene 257 (1), 119-130 (2000)

JOURNAL

MEDLINE

20510030

PUBMED

11054574

2 (bases 1 to 230000)

REFERENCE

AUTHORS

Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,

Moss, P., Paepel, B. and Wang, K.

TITLE

Direct Submission

Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,

Bothell, WA 98021, USA

FEATURES

Location/Qualifiers

1. 230000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="19"

/map="19q13"

/complement(<6963..>11517)

/gene="KLK1"

complement(join(<6963..7118,7668..7804,7923..8212,

9482..9641,11472..>11517))

/gene="KLK1"

/product="renal kallikrein"

complement(join(6963..7118,7668..7804,7923..8212,

9482..9641,11472..11517))

/gene="KLK1"

/note="serine protease"

/codon_start=1

/product="renal kallikrein"

/protein_id="AAG33353.1"

/db_xref="GI:11244758"

/translation="MWPLVCLALSLSGGTGAAPFIQSRIVGVGMECEHSPQWQALVH

FSTFGCGILVHRQWVLTAAHCTSDNYQLMGRHNLFDDETAFVHVSFPHGGEN

MSLLENHTRQADEYSHDLMLSLTEPADTITDAVKVELPTEPEVSGTCLASGWS

IEFENSFDDLCQVDKILPNBECKKAHVQVTDFMLCVHLEGGKDTTCVDSGGPL

MCDGVQGVTSWGVPGCTGPKPSVAVRVLSYVKVIEDTIAENS"

complement(join(<13552..13704,14377..14530,14678..14917,

15416..15569,19204..>19246))

/product="ACO protease"

complement(join(13552..13704,14377..14530,14678..14917,

15416..15569,19204..19246))

/note="serine protease"

/codon_start=1

/product="ACO protease"

/protein_id="AAG33354.1"

/db_xref="GI:11244759"

/translation="MWLLTLFLSLAATAQDGKLEGECAHSPQWQALVERGR

FNCASLISPHWLVSAAHCQSRFMRLGHEHNLKRDGPQLRTTSRVIPHPYEARS

HRNDILRLVQPARLNPOVRPAVLPTRCPHGEACVSVGMLVPLSPVSLPDLHC

ANTISIISDTSCKSYPCGLTNTMVCAGAEGRGAESCEGDSGGPLVCGILQIVSWG

VPCDNTKPGVYTKVCHYLEWIRETKEN"

<42595..>47769

/gene="KLK3"

join(<42595..42640,43880..44039,45669..45955,46099..46235,

47614..>47769)

/gene="KLK3"

/product="prostate specific antigen"

join(42595..42640,43880..44039,45669..45955,46099..46235,

47614..47769)

/gene="KLK3"

/note="serine protease"

/codon_start=1

/product="prostate specific antigen"

/protein_id="AAG33355.1"

/db_xref="GI:11244760"

/translation="MWVTVVFLTSLVTWIGAAPLILSRIVGVGMECEHSPQWQVLVAS

RGAVCGGVLPQWVLTAAHCTRNKSVILLGRSLPHPEDTGVQFVSHSPHPLVAD

MSLLKNFLRPGDSSHDMLRLSEPAELTDAVKMDLPTQEBALGTTTCVASCWSGI

EPFELTPKKLQCVDLHIVSDCAQVHPQKVTKFMLCAGRWTKGKTCSDSGSGLV

CNGVLOQITSWGSEPCALPERPSLYTKVHYRWIKDTIVANP"

<61139..>66229

/gene="KLK2"

join(<61139..61184,62391..62550,64142..64428,64542..64678,

66074..>66229)

/gene="KLK2"

```
CDS
/product="glandular kallikrein"
/join(61139..61184,62391..62550,64142..64428,64542..64678,
66074..66229)
/gene="KLK2"
/notes="serine protease"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MWDLVLSIALSVGTGAVPLIQSRIVGVGCEKHSOPWQVAVYS
HGWACGGVLVHPQWLTAACHLKNQVWLGRNLPEPDTGQRPVSHSFPHPLYN
ESLKHQSLRDESDHLLMLRLSEPAKITDVVKVLGLPTQEPALGTCTYASGWSGI
EPFEFLPRSLQCVSLHLLNDMCARAYSEKVFTEFMCAGLWGTGKTCGDSGGPLV
CNGVLOGITSPPFCALPEKPAVYTVVHYRWIKDTIAANP"
complement(join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..98405))
/product="serine protease"
/complement(join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPGWFLGVLILVAGSLVSGSCQIINGEDCSPHSQPW
QAALVMENELFCSGVLVHPQWLTAACHLKNQVWLGRNLPEPDTGQRPVSHSFPHPLYN
VRPEYNRLPLANDMLIKLDESVESEDITRSIASQCPTAGNSCLVSGWGLIANGR
MPTVLQCVNVSWESEVCSKLYDPLHYPSMFCAGGHDQKDCNCGDSGGPLICNGYLQ
GLVSFGKAPCGQGVPGVYTNLCFTIEWIKTVQAS"
complement(<131301..>140303)
/gene="KLK5"
/notes="synonym: SCTE"
complement(join(<131301..131456,136310..136443,
136529..136785,137525..137690,140255..>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement(join(131301..131456,136310..136443,
136529..136785,137525..137690,140255..140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWLWLCALITALLGVTDARSDDSSRIINGSDDCMHTQPQAAL
LLRNLQYCGAVLVPQWLTAACHKRVFRVLGHYLSLSPVYESQOMFGVKSIPH
PGYHPGNSDLMLIKNRRIRPKDVRPNVSHCPASGATKCLVSGWGTKSPQVHF
PKVLQCLNISVLQKCEDAYPRQIDDTMFCAGRKAGRDCSQGDSGSPVVCNGLSLQ
VSGDYPCARNRFGVYNLCKFTKWIQETIQANS"
complement(join(<146834..146986,149628..149764,
151186..151433,155052..155208,155948..>155987))
/product="protease M"
complement(join(146834..146986,149628..149764,
151186..151433,155052..155208,155948..155987))
/notes="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKLIMVYLSLIAAAWEQNKLVHGGPCDKTSHPYQALYTSGH
LLCGVLHPDWLWTAACHKPNLQVFLGRNLKQRESSQSQSVRAVHPDFAAS
HDQIMLLRLARPAKLSLQPLERDCSANTTSCHILGWGTADGFPDPTQDAYI
HLVRECEHAYPQGITQNMLCAGDERKYGKDCSQGDSGGPLVCGDHLRGLVSWGNI
GSKEKPGVYTVNVCRYTNWIKTIQAK"
complement(<165420..>170283)
/gene="KLK7"
/notes="synonym: SCCE"
complement(join(<165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement(join(165420..165575,167672..167808,
168124..168371,169651..169798,170211..170283))
/gene="KLK7"
/notes="serine protease"
gene
1 GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGATACCGACCCATCTCTGC 60
Db 136548 GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGATACCGACCCATCTCTGC 136489
Qy 61 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTCTGTTTGACAGTGCACCTTGCCTAAGG 120
Db 136488 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTCTGTTTGACAGTGCACCTTGCCTAAGG 136429
Qy 121 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAAGTGCAGAGATGCTTACCCGA 180
Db 136428 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAAGTGCAGAGATGCTTACCCGA 136369
Qy 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTCTCTGCCAGG 240
Db 136368 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTCTCTGCCAGG 136309
Qy 241 TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCACACTCGGTAAACATGGAGC 300
Db 136308 TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCACACTCGGTAAACATGGAGC 136249
Qy 301 GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGCAGTTCAGGACCCCTGTTTTCACAGA 360
Db 136248 GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGCAGTTCAGGACCCCTGTTTTCACAGA 136189
Qy 361 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGCTGCTTATGTCATTTC 420
Db 136188 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGCTGCTTATGTCATTTC 136129
Qy 421 CAGTTTTTTAGTAGCCACATTTAAAAACAGGTAAAAAAGGCTGGGCGCAGTGCTCACACCT 480
```

```
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MARSLLPLQILLLSLALETAGEAAGDKIIDGAPCARSGHPWQ
VALISQNLHCGGVLVNERWLTAACHKNMEYTVHLGSDTLGDRRAQRIKAKSKFRHP
SDLMQYHNDMLVNLKNSQARLSMVKVRLPSRCEPPGTTCTVSGWGTTSPDVTTP
SYLTCVDVKLSIPQDCTKVYKDLLENMLCAGIPDSKKNACNGDSGGPLVCRGTLQGL
VSMGTFCGPNPDGVTYQVCKFTWINDTKKXHR"
complement(join(<183943..184098,185635..185768,
187865..188127,188293..188452,188967..>189036))
/product="neuropsin"
complement(join(183943..184098,185635..185768,
187865..188127,188293..188452,188967..189036))
/notes="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRPRPRAAKTMFWLLLLGGAWAGHRAQBDKVLGGHECOPHSQ
PWQALPOGOQLLCGGVLVGNWLVTAACHKPKYTVRLGDHSLQNDKGPEOIEPVVQ
SIHPCVNSSDVEDHNDMLLOLRDOASIGSKVKPISLADHCTQPCQKCTVSGWGT
TSRENPPDLNCAEVKIFQKCEDAYPQGITDMVCAGSSKAGATCGQDSGGPLVC
DGAQGITSGSDPCGRSDRPGVYTNICRILDMIKIISRGK"
complement(join(<190980..191129,191573..191709,
194324..194589,197048..197204,197370..>197412))
/product="kallikrein-like 3"
complement(join(190980..191129,191573..191709,
194324..194589,197048..197204,197370..197412))
/notes="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translation="MKLGLCALLSLLAGHWADTRAIGAEBCPNQWQAGLPHLT
RLFCGATLISDRMLLTAACHKPKYLVRLGEHLHWKEGPEQLFRVTDFFPFGPKND
LSANDHNDIMLRLPRQARLSPAVQINLSQTCVSPGMQLCSLWGNVSPKALFV
TLQCANISILENKLCHWAYPSGHI SDSMLCAGLWEGGRGSCQDGGSGPLVNCNTLAGVV
SGGAEPSCRPRRPVAVTVSVCHYLDWIQEIWEN"
```

```
Query Match 100.0%; Score 1001; DB 9; Length 2300000;
Best Local Similarity 100.0%; Pred. No. 1.3e-227;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGATACCGACCCATCTCTGC 60
Db 136548 GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGATACCGACCCATCTCTGC 136489
Qy 61 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTCTGTTTGACAGTGCACCTTGCCTAAGG 120
Db 136488 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTCTGTTTGACAGTGCACCTTGCCTAAGG 136429
Qy 121 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAAGTGCAGAGATGCTTACCCGA 180
Db 136428 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAAGTGCAGAGATGCTTACCCGA 136369
Qy 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTCTCTGCCAGG 240
Db 136368 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTCTCTGCCAGG 136309
Qy 241 TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCACACTCGGTAAACATGGAGC 300
Db 136308 TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCACACTCGGTAAACATGGAGC 136249
Qy 301 GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGCAGTTCAGGACCCCTGTTTTCACAGA 360
Db 136248 GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGCAGTTCAGGACCCCTGTTTTCACAGA 136189
Qy 361 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGCTGCTTATGTCATTTC 420
Db 136188 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGCTGCTTATGTCATTTC 136129
Qy 421 CAGTTTTTTAGTAGCCACATTTAAAAACAGGTAAAAAAGGCTGGGCGCAGTGCTCACACCT 480
```


* 17511 17610: gap of 100 bp
 * 17611 24602: contig of 6992 bp in length
 * 24603 24702: gap of 100 bp
 * 24703 35434: contig of 10732 bp in length
 * 35435 35535 124474: gap of 100 bp
 * 35535 124474: contig of 8940 bp in length
 * 124475 124574: gap of 100 bp
 * 124575 134664: contig of 10090 bp in length
 * 134665 134765 162343: contig of 27579 bp in length
 * 134765 162343: contig of 27579 bp in length
 * 162344 162443: gap of 100 bp
 * 162444 208917: contig of 46474 bp in length
 * 208918 209017: gap of 100 bp
 * 209018 217346: contig of 8329 bp in length.

FEATURES

Location/Qualifiers
 1. 217346
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19"
 /clone="RP11-795B6"
 /clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. 8149
 /note="assembly_fragment"
 clone end:SP6
 vector side:left

misc_feature

8250..9592
 /note="assembly_fragment"

misc_feature

9693..10733
 /note="assembly_fragment"

misc_feature

10834..13519
 /note="assembly_fragment"

misc_feature

13620..17510
 /note="assembly_fragment"

misc_feature

17611..24602
 /note="assembly_fragment"

misc_feature

24703..35434
 /note="assembly_fragment"

misc_feature

35535..124474
 /note="assembly_fragment"

misc_feature

124575..134664
 /note="assembly_fragment"

misc_feature

134765..162343
 /note="assembly_fragment"

misc_feature

162444..208917
 /note="assembly_fragment"

misc_feature

209018..217346
 /note="assembly_fragment"
 clone end:T7
 vector side:right

ORIGIN

Query Match 98.3%; Score 984.2; DB 2; Length 217346;
 Best Local Similarity 99.6%; Pred. No. 1.3e-223;
 Matches 997; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 Qy 1 GACACCAAGAGCCCCCAGGTGAGTGCAGGTTCTTGTATACCGACCCATCTCTGC 60
 Db 198410 GACACCAAGAGCCCCCAGGTGAGTGCAGGTTCTTGTATACCGACCCATCTCTGC 198351
 Qy 61 CGCCTTCATCTTTCTCCACTTCTCATTTGTTCTGTTTGCAGTGCACATTCCTTAAG 120
 Db 198350 CGCCTTCATCTTTCTCCACTTCTCATTTGTTCTGTTTGCAGTGCACATTCCTTAAG 198291
 Qy 121 TCCTCCAGTGTGTAATATCAGCGTGTAAAGTGCAGAAAAGTGCAGGATGCTTACCCGA 180
 Db 198290 TCCTCCAGTGTGTAATATCAGCGTGTAAAGTGCAGAAAAGTGCAGGATGCTTACCCGA 198231
 Qy 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGATCTCTGCCAGG 240
 Db 198230 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGATCTCTGCCAGG 198171

Qy 241 TGAGGACACCTCTCTTTATTTCAGAGATACACACTGAGTGCCCACTCGGTAAACATGAGC 300
 Db 198170 TGAGGACACCTCTCTTTATTTCAGAGATACACACTGAGTGCCCACTCGGTAAACATGAGC 198111
 Qy 301 GTTGCCCAAAATCTGAGAAATCCAGCAATTCGCAAGACAGTCCAGGACCCCTGTTCTCAGCA 360
 Db 198110 GTTGCCCAAAATCTGAGAAATCCAGCAATTCGCAAGACAGTCCAGGACCCCTGTTCTCAGCA 198051
 Qy 361 GCTCATACCCCTAGAGTAGTGGTGTATTAGTAAATAATAGTGCAGTCTCTTATGTGTCATTTC 420
 Db 198050 GCTCATACCCCTAGAGTAGTGGTGTATTAGTAAATAATAGTGCAGTCTCTTATGTGTCATTTC 197991
 Qy 421 CAGTTTTTTAGTAGCCACATTAATAACAGGTAAATAAGGCTGGGCGCAGTGGCTCACACCT 480
 Db 197990 CAGTTTTTTAGTAGCCACATTAATAACAGGTAAATAAGGCTGGGCGCAGTGGCTCACACCT 197931
 Qy 481 GTAATCCAGCAGCTTTGGGAGGCTGAGGAGGAGATCAGCTTTGGTCAGGAGTTTCAGA 540
 Db 197930 GTAATCCAGCAGCTTTGGGAGGCTGAGGAGGAGATCAGCTTTGGTCAGGAGTTTCAGA 197871
 Qy 541 CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAAAATACAAAAAATTAGCCTGGC 600
 Db 197870 CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAAAATACAAAAAATTAGCCTGGC 197811
 Qy 601 ATGGTGGCGGCGCCTGTAAATCTCAGCTGCTCAGGAGGCGGAGACACAAGAAATCACTTAA 660
 Db 197810 ATGGTGGCGGCGCCTGTAAATCTCAGCTGCTCAGGAGGCGGAGACACAAGAAATCACTTAA 197751
 Qy 661 ACCGAGGAGTGGAGGTTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 720
 Db 197750 ACCGAGGAGTGGAGGTTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 197691
 Qy 721 AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG 780
 Db 197690 AGAGTGACACTTTTGTCTCAAAAAG-AAAAAACAAGTAAAAAAGAAACAGGTGAAG 197632
 Qy 781 TTAACCTTTAATAAACCCCAATGTATCCCAATACAAATCAATTTCAAAGTGTAAATATAATAA 840
 Db 197631 TTAACCTTTAATAAACCCCAATGTATCCCAATACAAATCAATTTCAAAGTGTAAATATAATAA 197572
 Qy 841 ACAATTATGAATGAGATACATTTACATTTCTTTCTTTGTTTTCATATTAAAGTCTTGAAGT 900
 Db 197571 ACAATTATGAATGAGATACATTTACATTTCTTTCTTTGTTTTCATATTAAAGTCTTGAAGT 197512
 Qy 901 GAGTATATATGTTATGCTGCAGCAGCATCTCAATTTGGAGTGCAGTGCAGTGCAGTGCCTC 960
 Db 197511 GAGTATATATGTTATGCTGCAGCAGCATCTCAATTTGGAGTGCAGTGCAGTGCAGTGCCTC 197452
 Qy 961 AGTAGCCACATGTGGCTAGCAGTTACTGTATGGATGGCAC 1001
 Db 197451 AGTAGCCACATGTGGCTAGCAGTTACTGTATGGATGGCAC 197411

RESULT 8

AC130782/c

LOCUS

DEFINITION

AC130782

Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12

ordered pieces.

AC130782

VERSION

AC130782.2 GI:25167101

KEYWORDS

HTG; HTGS PHASE2; HTGS DRAFT.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

1 (bases 1 to 200792)

AUTHORS

Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Loric, P., Lee-Lin, S.-Q., Lega, P., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C., Maskeri, B., McDownell, A., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,

Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stanton, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 200792)
Green, E.D.
Direct Submission
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 200792)
Green, E.D.
Direct Submission
Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On Nov 22, 2002 this sequence version replaced gi:22218452.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: dhz
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
1 28306: contig of 28306 bp in length
* 28307 28406: gap of unknown length
* 28407 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length
* 37957 73522: contig of 35566 bp in length
* 73523 73622: gap of unknown length
* 73623 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 88817: contig of 5150 bp in length
* 88818 88917: gap of unknown length
* 88918 125611: contig of 36694 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length

* 174699 174798: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186483: gap of unknown length
* 186483 193344: contig of 6862 bp in length
* 193345 193445: gap of unknown length
* 193445 193663: contig of 5919 bp in length
* 193664 194663: gap of unknown length
* 194664 200792: contig of 1329 bp in length.
FEATURES
source
location/Qualifiers
1..200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"
1..28306
/note="assembly_fragment"
clone_end:SP6
vector_side:left
28407..37856
/note="assembly_fragment"
37957..73522
/note="assembly_fragment"
73623..83567
/note="assembly_fragment"
83668..88817
/note="assembly_fragment"
88918..125611
/note="assembly_fragment"
125712..159879
/note="assembly_fragment"
159980..174698
/note="assembly_fragment"
174799..186382
/note="assembly_fragment"
186483..193344
/note="assembly_fragment"
193445..199363
/note="assembly_fragment"
199464..200792
/note="assembly_fragment"
clone_end:T7
vector_side:right
ORIGIN
Query Match 95.8%; Score 958.8; DB 2; Length 200792;
Best Local Similarity 98.6%; Pred. No. 1.5e-217;
Matches 988; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
Qy 1 GACACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGTATACCGACCATCTCTGC 60
Db 73157 GACACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGTATACCGACCATCTCTGC 73098
Qy 61 CGCCTTCACCTTCTCCACTTCTCATTTGTTCTCTTTGACAGTGCACCTTCCCTAAGG 120
Db 73097 CGCCTTCACCTTCTCCACTTCTCATTTGTTCTCTTTGACAGTGCCTTCCCTAAGG 73038
Qy 121 TCCTCCAGTGTCTGAATATCAGCGTGTAAAGTGCAGAAAGTGCAGGATGCTTACCCGA 180
Db 73037 TCCTCCAGTGTCTGAATATCAGCGTGTAAAGTGCAGAAAGTGCAGGATGCTTACCCGA 72978
Qy 181 GACAGATAGATGACACCATGTTTTCGCCCGGTGACAAAGCAGGTAGAGATCTCTGCCAAG 240
Db 72977 GACAGATAGATGACACCATGTTTTCGCCCGGTGACAAAGCAGGTAGAGATCTCTGCCAAG 72918
Qy 241 TGAGGACACCTCTCTTTATTTCAGCAGATACACATGAGTGCACACTCGTAACTGGAGC 300
Db 72917 TGAGGACACCTCTCTTTATTTCAGCAGATACACATGAGTGCACACTCGTAACTGGAGC 72858
Qy 301 GTTCCCAAATTTCTGAGAAATCCAGCAATTCGCAAGACAGTCAAGGACCCCTGTTCTCACAGA 360
Db 72857 GTTCCCAAATTTCTGAGAAATCCAGCAATTCGCAAGACAGTCAAGGACCCCTGTTCTCACAGA 72798

```
Qy 361 GCTCATACCCCTAGAGTAGTGTGTTTAGTAGAATAATGCTGAGCTGCTTATGTCATTTC 420
Db |||||||
Qy 72797 GCTCATACCCCTAGAGTAGTGTGTTTAGTAGAATAATGCTGAGCTGCTTATGTCATTTC 72738
Db |||||||
Qy 421 CAGTTTTTTTAGTAGCCCATTTAAACAGGTAAACAGGCTGGGCGCAGTGGCTCACACCT 480
Db |||||||
Qy 481 GTAATCCAGCAGCTTTGGGAGGCTGAGCAGGCGAGATCACTTTGGTCAGGAGTTTGAGA 540
Db |||||||
Qy 72677 GTAATCCAGCAGCTTTGGGAGGCTGAGCAGGCGAGATCACTTTGGTCAGGAGTTTGAGA 72618
Db |||||||
Qy 541 CTAGCTGGCCAGCATGGCGAACTCTGTCTCTAAAAAATAACAAAAATAGACCTGGC 600
Db |||||||
Qy 72617 CCAGCTGGCCAGCATGGCGAACTCTGTCTCTAAAAAATAACAAAAATAGACCTGGC 72558
Db |||||||
Qy 601 ATGCTGGCGGCGCTCTGTAATCTCAGCTGCTCAGGAGCGGAGACACAAGATCACTTAA 660
Db |||||||
Qy 661 ACCAGAGGTGGAGGTTGCACTGAGCTGAGATCGTGCCACT-CACTCCACCTGGGAGA 719
Db |||||||
Qy 72497 ACCAGAGGTGGAGGTTGCACTGAGCTGAGATCGTGCCACTGACCTCCACCTGGGAGA 72438
Db |||||||
Qy 720 CAGAGTCACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAAGAACAGGTGAA 779
Db |||||||
Qy 72437 CAGAGTCACACTTTTGTCTCAAAAAAG-AAAAAACAAGTAAAAAAAGAACAGGTGAA 72379
Db |||||||
Qy 780 GTTAACCTTTAATAACCCCAATGATCCCAATACAAATCATTTCAAAGTGTAAATTAATAA 839
Db |||||||
Qy 840 AACAAATATGAATAGATACCTTTACATCTTTCTTTTTCATATTAAGTCCTTTGAAG 899
Db |||||||
Qy 72318 AACAAATATGAATAGATACCTTTACATCTTTCTTTTTCATATTAAGTCCTTTGAAG 72259
Db |||||||
Qy 900 TGAGTATATATGTTATGCTGACAGACATCTCAATTTGGACTAGCTACATTTCAAGTGCT 959
Db |||||||
Qy 72258 TGAGTATATATTTATGCTGACAGACATCTCAATTTGGACTAGCTACATTTCAAGTGCT 72199
Db |||||||
Qy 960 CAGTAGCCACATGCGGCTAGCAGTACTGTATTTGGATGGAC 1001
Db |||||||
Qy 72198 CAGTAGCCACATGCGGCTAGCAGTACTGTGTGTGATGGAC 72157
Db |||||||

RESULT 9
AC098800
LOCUS Homo sapiens BAC clone RP11-666P13 from 2, complete sequence.
DEFINITION AC098800 AC068744
ACCESSION AC098800.3 GI:23130726
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 142003)
Toward a complete human genome sequence
Sulston, J.E. and Waterston, R.
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 142003)
AUTHORS Shahid, S., Meyer, R. and Spalding, L.
TITLE The sequence of Homo sapiens BAC clone RP11-666P13
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 142003)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 142003)
```

AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Waterston, R.H.
Direct Submission
Submitted (08-AUG-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 142003)
Waterston, R.H.
Direct Submission
Submitted (18-SEP-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 142003)
Waterston, R.
Direct Submission
Submitted (03-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 142003)
Waterston, R.
Direct Submission
Submitted (15-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 18, 2002 this sequence version replaced gi:22002231.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0666P13
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPC111 human BAC library was made from the blood of one male
donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
This sequence is not the entire insert of the clone. This clone is
overlapped by AC097532 and AC104405.

Uncertain bases from base 9817 to base 9869.

There is an unresolved tandem repeat from base 9870 to 11201.

Data from AC068744 and AC104405 was used to finish this clone.

The sequence of AC068744 has been incorporated into AC098800.
Location/Qualifiers

FEATURES


```

RESULT 10
AC133961/c
LOCUS      AC133961      177787 bp      DNA      linear      PRI 03-JAN-2003
DEFINITION Homo sapiens BAC clone RP13-494C23 from 4, complete sequence.
ACCESSION  AC133961
VERSION     AC133961.3  GI:26190601
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 177787)
            Sulston, J.E. and Waterston, R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE    99063792
PUBMED     9847074
REFERENCE  2 (bases 1 to 177787)
            Isak, A., Bielicki, L., Creason, K. and Cotton, M.
            The sequence of Homo sapiens BAC clone RP13-494C23
            Unpublished (2001)
REFERENCE  3 (bases 1 to 177787)
            Waterston, R.H.
            Direct Submission
            Submitted (20-SEP-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  4 (bases 1 to 177787)
            Waterston, R.H.
            Direct Submission
            Submitted (05-NOV-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  5 (bases 1 to 177787)
            Waterston, R.H.
            Direct Submission
            Submitted (07-DEC-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  6 (bases 1 to 177787)
            Waterston, R.
            Direct Submission
            Submitted (03-JAN-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Dec 7, 2002 this sequence version replaced gi:24580489.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_FH0494C23
            -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-13 Human Female BAC Library was constructed using improved cloning techniques developed by Kazutoyo Osoegawa. The library was generated by Baohui Zhao in our laboratory. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01HG01165-03). The library was generated according to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA Sequencing'.

Female blood was obtained via a double-blind selection protocol. Female blood DNA was isolated from one randomly chosen donor (out of 10 female donors) and partially digested with a combination of EcoRI and EcoRI Methylase for library segments 1&2 or either MboI or DpnII for library segments 3&4. Size selected DNA was cloned into the pBACe3.6 vector between the EcoRI sites for library segments 1&2 or the BamHI sites for library segments 3&4. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). The library has been arrayed into 384-well microtiter dishes and also gridded onto 22x22cm nylon high density filters for screening by probe hybridization.

NEIGHBORING SEQUENCE INFORMATION:

The sequence from base 67288 to base 67340 was derived from one plasmid subclone.

Polymorphisms have been identified between AC020706 and this sequence.

Data from AC020706 was used to finish this clone.

This sequence is not the entire insert of the clone. This clone is overlapped by AC020706 and AC093660.

FEATURES	source
Location/Qualifiers	1..177787
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone_lib="RPCI-13"
	/clone="RP13-494C23"
repeat_region	4..70
	/rpt_family="MIR"
repeat_region	703..833
	/rpt_family="Alu"
repeat_region	852..875
	/rpt_family="AT-rich"
repeat_region	911..1138
	/rpt_family="Alu"
repeat_region	1157..1191
	/rpt_family="(CAAAA)n"
repeat_region	1341..1408
	/rpt_family="GA-rich"
repeat_region	1432..1577
	/rpt_family="GA-rich"
repeat_region	2115..2458
	/rpt_family="L2"
repeat_region	2562..2603
	/rpt_family="MER1_type"
repeat_region	2773..2868
	/rpt_family="MER1_type"
repeat_region	2932..3240
	/rpt_family="Alu"
repeat_region	3264..3462
	/rpt_family="MaLR"
repeat_region	3463..3760
	/rpt_family="Alu"
repeat_region	3761..3934
	/rpt_family="MaLR"
repeat_region	5508..5796
	/rpt_family="Alu"
repeat_region	6977..7240
	/rpt_family="Alu"

[illegible]

on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>
 RPI-71H19 is from the library RPCI-1 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pCYPAC2

This sequence is the entire insert of clone RPI-71H19 The true left
 end of clone RPI-216J23 is at 87497 in this sequence. The true
 right end of clone RPI-422B11 is at 29670 in this sequence.

FEATURES

source

```

1..138411
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RPI-71H19"
/clone_lib="RPCI-1"
complement(477..1095)
/notes="match: GSS: Em:AQ506907"
complement(609..1046)
/notes="match: GSS: Em:AQ437147"
complement(620..1089)
/notes="match: GSS: Em:AQ819330"
complement(621..1083)
/notes="match: GSS: Em:AQ437147"
complement(761..1000)
/notes="match: GSS: Em:B71825"
complement(977..1090)
/notes="match: GSS: Em:AQ552044"
1514..1858
/notes="match: GSS: Em:AQ176950"
1545..1851
/notes="match: GSS: Em:AQ176950"
1852..2161
/notes="AluXq repeat: matches 1..309 of consensus"
2510..8661
/notes="L1PFA repeat: matches 13..6142 of consensus"
8448..8976
/notes="match: GSS: Em:AQ347835"
8662..8847
/notes="match: GSS: Em:AQ347835"
8848..9156
/notes="MLTII repeat: matches 40..410 of consensus"
9172..9481
/notes="AluSg repeat: matches 5..313 of consensus"
9692..9804
/notes="MLTII repeat: matches 267..371 of consensus"
9824..10043
/notes="MIR repeat: matches 18..251 of consensus"
10338..10689
/notes="MLT2PA repeat: matches 9..356 of consensus"
complement(10977..11390)
/notes="match: STS: Em:G42414"
complement(11825..12306)
/notes="match: GSS: Em:AQ141424"
12402..12825
/notes="match: GSS: Em:AQ266412"
12647..12722
/notes="MIR repeat: matches 1..79 of consensus"
13978..14287
/notes="AluDb repeat: matches 1..310 of consensus"
14664..15035
/notes="L2 repeat: matches 2294..2673 of consensus"
15036..15342
/notes="AluSg repeat: matches 1..308 of consensus"
15343..15421
/notes="L2 repeat: matches 2673..2750 of consensus"
15485..15534
/notes="MIR repeat: matches 213..262 of consensus"

```

```

15581..15703
/notes="MIR repeat: matches 110..234 of consensus"
16310..16535
/notes="MIR repeat: matches 21..231 of consensus"
16597..17598
/notes="L1P7 repeat: matches 5135..6145 of consensus"
17617..17827
/notes="AluSp repeat: matches 83..309 of consensus"
17828..19349
/notes="L1P7 repeat: matches 3624..5146 of consensus"
complement(19723..20194)
/notes="match: GSS: Em:A2016849"
20519..20665
/notes="match: GSS: Em:AQ371979"
20807..21136
/notes="match: GSS: Em:AQ508843"
20807..21096
/notes="match: GSS: Em:AQ508843"
20818..21272
/notes="match: GSS: Em:AQ209413"
20818..21096
/notes="match: GSS: Em:AQ209413"
20831..21139
/notes="match: GSS: Em:AQ514718"
20902..21139
/notes="match: GSS: Em:AQ093115"
21097..21176
/notes="40 copies 2 mer ct 76% conserved"
21524..22065
/notes="match: GSS: Em:AQ530328"
22856..23094
/notes="AluSg repeat: matches 1..239 of consensus"
23102..23135
/notes="17 copies 2 mer aa 88% conserved"
24117..24204
/notes="MER5B repeat: matches 1..90 of consensus"
24883..25185
/notes="AluY repeat: matches 1..300 of consensus"
25364..25449
/notes="MSTD repeat: matches 1..96 of consensus"
25932..26300
/notes="L1MB8 repeat: matches 5779..6162 of consensus"
26483..26600
/notes="L2 repeat: matches 2583..2703 of consensus"
27463..27633
/notes="AluSp repeat: matches 137..308 of consensus"
complement(27586..28035)
/notes="match: GSS: Em:AQ753792"
complement(27650..27736)
/notes="match: GSS: Em:AQ753792"
27737..28017
/notes="AluJ2 repeat: matches 1..276 of consensus"
28766..28860
/notes="MIR repeat: matches 85..188 of consensus"
29043..29225
/notes="FRAM repeat: matches 0..174 of consensus"
29387..29497
/notes="MIR repeat: matches 29..143 of consensus"
29585..29879
/notes="AluSp repeat: matches 1..296 of consensus"
31001..31343
/notes="L1MD2 repeat: matches 5957..6332 of consensus"
31344..31643
/notes="AluSc repeat: matches 1..301 of consensus"
31644..31712
/notes="L1MD2 repeat: matches 5890..5957 of consensus"
31714..31863
/notes="L1MD3 repeat: matches 7589..7735 of consensus"
31874..32598
/notes="L1MD2 repeat: matches 5386..6122 of consensus"
33641..33934
/notes="AluX repeat: matches 3..298 of consensus"
33933..33941

```

```
/note="single clone region"
34941..35324
/note="match: GSS: Em:B55255"
35189..35758
/note="L2 repeat: matches 2189..2747 of consensus"
complement(36539..37046)
/note="match: GSS: Em:AQ58962"
36809..37468
/note="match: GSS: Em:AQ541391"
36811..37282
/note="match: GSS: Em:AQ421653"
36831..37467
/note="match: GSS: Em:AQ547524"
36992..37051
/note="MER95 repeat: matches 86..139 of consensus"
37054..37177
/note="MER95 repeat: matches 1..139 of consensus"
37956..38473
/note="match: GSS: Em:AQ246505"
complement(37968..38506)
/note="match: GSS: Em:AQ427574"

Query Match 27.4%; Score 274.4; DB 9; Length 138411;
Best Local Similarity 78.0%; Pred. No. 1.3e-54;
Matches 415; Conservative 0; Mismatches 91; Indels 26; Gaps 6;

QY 472 CTCACACCTGTAATCCAGCACTTTGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAGG 531
Db 43069 CTCACACCTGTAATCTCAGCACTTTGGAGGCGGAGGCGGAGATCACTTTGGTCAGG 43128
QY 532 AGTTTGAGACTAGCTCGCCAACTGCGCAACTCTGCTCTTAAAAAATAACAATAAT 591
Db 43129 AGTTTGAGACTAGCTCGCCAACTGCGCAACTCTGCTCTTAAAAAATAACAATAAT 43187
QY 592 TAGCTGTCATGTTGGCGGCGGCGCTGTAATCTCAGCTGCTCAGGAGGCGGAGGAGCAACA 651
Db 43188 TAGCAGGCGCATGTTGGTGGTACCTGTAATCCAGCTACTCAGGAGGCTGAGGAGGAGA 43247
QY 652 ATCACTTAACCCAGGAGTGAGGTTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 710
Db 43248 ATTGCTTAACCCAGGAGGAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 43307
QY 711 CCTGGGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAAGAAAAAGTAAAAAGAA 770
Db 43308 CTGGGTGACAGAGGAGACTCTCTCTAAAAAAGAAAAAAGAAAAAGTAAAAAGAA 43367
QY 771 ACAGG-----TGAAGTTAACTTTTAATACCAATGTATCCCAATACAATCATTTTC 821
Db 43368 GGATGAAGTTAATTTTAAATTTTATTTGATTTAACTTAATTTGCAATATTTGATTTT 43427
QY 822 AAAGTGAATTAATAAACAATTAATGAATGAGATATTTACATCTTTTCTTTGTTTTC 881
Db 43428 GATATATAATTTAGTATACAAAGTT-----ATTGAGATATTTTACATCTTTTCTCT 43478
QY 882 ATATTAAGCTTTGAAAGTGAATATATGTTATGCTGACGACATCTCAATTTGGACT 941
Db 43479 ATGTAAGCTTTGAAATTTTGTGATATTTTATG-----GTATATCTCAATTTGGACT 43532
QY 942 AGCTACATTTTCAGGTGCTCAGTAGCCATGTTGGCTAGCAGTTACTGTTATTG 993
Db 43533 AGCCACATTTCAAGAACTCAGTAGCCATGTTGGCTAGTACTGTTATTG 43584

RESULT 12
AL355815
LOCUS
DEFINITION
Human DNA sequence from clone RP3-453H5 on chromosome 6. Contains
ESTs, STSs, GSSs and a CpG island. Contains a novel gene, complete
sequence.
ACCESSION
AL355815
VERSION
AL355815.9 GI:10443432
KEYWORDS
HTG; CpG island.
SOURCE
Homo sapiens (human)
```

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 92882)

REFERENCE

AUTHORS

TITLE

JOURNAL

Parker, A.

Direct Submission

Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

COMMENT

On Oct 1, 2000 this sequence version replaced gi:10186513.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/HGP/Chr6
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6

IMPORTANT: This sequence is not the entire insert of clone
RP3-453H5. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.

The true left end of clone RP3-453H5 is at 1 in this sequence. The
true left end of clone RP3-470K1 is at 92783 in this sequence. This
sequence has been finished according to sequence map criteria as
follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
ambiguous, there is an annotation using the 'unsure' feature key.
RP3-453H5 is from the library RPCI-3 constructed by the group of
Piet de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES

source

Location/Qualifiers

1..92882

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="6"

/clone="RP3-453H5"

/clone_lib="RPCI-3"

4..33

/note="Alu repeat: matches 26..55 of consensus"

29..566

/note="match: STS: Em:HS453H5T"

2269..2693

/note="L1ME2 repeat: matches 5523..5962 of consensus"

2764..2993

/note="L1ME2 repeat: matches 5141..5391 of consensus"

2994..3398

/note="L1MD3 repeat: matches 7338..7740 of consensus"

3399..3414

/note="L1ME2 repeat: matches 5128..5141 of consensus"

3673..4100

/note="L1 repeat: matches 4439..4840 of consensus"

4101..4399

/note="AluX repeat: matches 1..299 of consensus"

4400..4765

/note="L1 repeat: matches 4004..4439 of consensus"

4772..5034

/note="L1M4 repeat: matches 3188..3469 of consensus"

5035..5355

/note="AluJo repeat: matches 1..309 of consensus"

complement(5142..5568)

/note="match: GSS: Em:B71585"

complement(5243..5586)

/note="match: GSS: Em:B71584"

repeat_region

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region


```
Db 50597 AGGCAGAGGTTCCAGTGAGTGAGATCATGCCACTGCAGTCAGCGCTGGGCAACAGAGTG 50656
Qy 727 ACACCTTTGTCTCAAAAAAGAAAAAAGAAAAACAGTAAAGAAACAGCGTGAAGTTAACT 786
Db 50657 AGTCTCTCTCTCAAAAAAGAAAAAAGAAAAAGGTT----ATATTAATTTTAAATAATATATT 50712
Qy 787 TTAATAACCCCAATGATCCCAATACAACTATTTCAAGAGTGTAATTAATATAATAAACAATT 846
Db 50713 TATTTACTCAATATAACCAAAATATATATTTCAATGCAATCAATATAAATAA-AGTT 50771
Qy 847 ATGAATGAGATCTTTTACATCTTTTCTTTTTCATATTAAGTCTTTTGAAGTGAGTAT 906
Db 50772 GTTAATGAATACCTTTTAAATCTGTT----TTTTTATACAAATCTTGGAAATCCAGTGT 50827
Qy 907 ATATGTTATGTCAGACAGACATCTCAATTTGGACTAGCTACATTTTCAGCTGCTCAG 962
Db 50828 GTATTTACACTTCCAGCATATTTTGGTTTGGACCAAGCCACATTTTCAAGTGTCTCAG 50883

RESULT 13
HS227L5 Human DNA sequence from clone RP6-227L5 on chromosome Xp11.22-11.3,
LOCUS complete sequence.
DEFINITION
ACCESSION AL031585
VERSION AL031585.2 GI:21212870
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 85304)
Direct Submission
Submitted (21-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hunquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:3980350.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/Celegans/wormpep/
This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP6-227L5 is from the library RP6-6 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC4.
FEATURES
source
1. 85304
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RP6-227L5"
/db_xref="taxon:9606"
/chromosome="X"
/maps="p11.22-11.3"
/clone="RP6-227L5"
```

```
ORIGIN
Query Match 26.1%; Score 261.6; DB 9; Length 85304;
Best Local Similarity 70.1%; Pred. No. 1.4e-51;
Matches 443; Conservative 0; Mismatches 169; Indels 20; Gaps 6;
/clone_lib="RPCI-6"
Qy 373 GAGTAGTGGTGTCTAGTAGAATAATGCTGAGCTGCTTATGTCATTTCCAGTTTGTAGT 432
Db 81598 GAAGATTTCCTTTTGGAGGTTATTAATCAAAATTTCTAGTAGTACATTTTAAAAATG 81657
Qy 433 AGCCACATTAACACAGGTAA--AAAGCTGGCGCAGTGGCTCACACTGTAAATCCAG 490
Db 81658 TTAATTTTAAAAAGTAATTTTCAGGCTGGACATCGTGGCTCACACTGTATCTCAG 81717
Qy 491 CACTTTGGAGGCTGAGGCGAGGATCACCTTTGGTCAGGAGTTTGAGACTAGCTGGC 550
Db 81718 CATTTGGAGTGTCTGAGGCGAGTGCATCTTGGTGGGAGTTTCGAGACCGAGCTGGC 81777
Qy 551 CAACATGGCGAAGTCTGCTCTTAAAAAATAAATAAATAAATAGCTGGCATGTGGCG 610
Db 81778 CAACATGGTGAACCTCATCTCT-ACTAAAAATAACAAAAATTCAGGCGCATGTGG 81836
Qy 611 GCGCTGTAAATCTCAGCTGCTCAGGAGCGCAGACACAAAGAAATCACTTTAAACCCAGGAGG 670
Db 81837 GTGCTGTAGTCCAGCTACTCAGAGGCTGAGGAGGAGAAATGCTTTAACCCAGGAGG 81896
Qy 671 TGGAGGTTGCTGAGCTGAGATCGTCCACT-CACTCCAACTGGGAGACAGAGTGACA 729
Db 81897 CAGAGGTTGCTGAGTGGAGCAAGATCACACCTCCAGCTCAGGCAACAGAGTGAGA 81956
Qy 730 CTTTTGTCTCAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 783
Db 81957 CTCAGTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 82016
Qy 784 ACTTTAATACCCCAATGATCCCAATACATCAATTTCAAGTGTAAATTAATAAACA 843
Db 82017 ACTTATTAACTTAATATATCCCAATTAATTAACACTTCAACTGTCAATTAATTA 82074
Qy 844 ATTATGATGAGATACCTTTTACATCTTTCTTTTGTGTTTTCATTAATGCTTTGAAAGTGAG 903
Db 82075 AATATTGTGAATACTTTATA-----TTTTTTTCATACTAAGTCTTTGAAATCTGC 82126
Qy 904 TATATATGTTATGCTGAGCAGACATCTCAATTTGGACTAGCTACATTTCCAGGTGCTCAGT 963
Db 82127 TGTGTTATTTATCTTACCAACATCTCAATTTGAACTAGGCACATACCAATGCTCAGT 82186
Qy 964 AGCCACATGTGGCTAGCAGTTACTGTATTGGA 995
Db 82187 AGCCACATGTGATCAACGGCTATCATATTGAA 82218

RESULT 14
LOCUS AC087591 161014 bp DNA linear PRI 20-DEC-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-95M5 map 3p, complete
sequence.
ACCESSION AC087591
VERSION AC087591.2 GI:24796758
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161014)
Zhang,H., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C.,
Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q.,
Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N.,
Li,C., Li,C., Li,F., Li,J., Li,L., Li,S., Li,T., Li,Y.,
Liu,N., Liu,B., Liu,Y., Li,W., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q.,
Qi,X., Song,L., Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R.,
Wang,H., Wang,J., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
```

Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L.,
Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B.,
Zhu, N., Yu, J. and Yang, H.
Chromosome 3p genomic sequence

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 161014)
Zhang, H., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C.,
Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q.,
Gu, X., Guo, D., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C.,
Li, C., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,
Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y.,
Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,
Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X.,
Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and
Yang, H.

TITLE
JOURNAL

Submitted (12-JAN-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

REFERENCE
AUTHORS

3 (bases 1 to 161014)
Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H.,
Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z.,
He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F.,
Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y.,
Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y.,
Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,
Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.

TITLE
JOURNAL

Submitted (08-NOV-2002) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

REFERENCE
AUTHORS

4 (bases 1 to 161014)
Zhang, H., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C.,
Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q.,
Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N.,
Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y.,
Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q.,
Qi, X., Song, L., Song, S., Sun, M., Sun, Y., Tan, X., Tao, R.,
Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X.,
Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C.,
Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.

TITLE
JOURNAL

Submitted (20-DEC-2002) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

COMMENT

On Nov 8, 2002 this sequence version replaced gi:12084032.

-----Genome Center
Center: Beijing Center
Center code: Beijing
Website: <http://hgci.gtp.ac.cn>
<http://www.genomics.org.cn>
Contact: hgc@gtp.ac.cn
-----Project Information
Center project name: 1k project
Center clone name: RP11-95M5
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 55% of reads
Assembly: Phrap; version 0.990329
Consensus quality: 162662 bases at least Q40
Consensus quality: 163749 bases at least Q30
Consensus quality: 164140 bases at least Q20
Insert size: 161014; sum-of-contigs
Quality coverage: 9.63x in Q20 bases; sum-of-contigs

FEATURES
source
1. 161014
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3p"
/map="3p"
/clone="RP11-95M5"

ORIGIN
Query Match 26.0%; Score 260; DB 9; Length 161014;
Best Local Similarity 67.7%; Pred. No. 3.7e-51;
Matches 447; Conservative 0; Mismatches 190; Indels 23; Gaps 5;
344 ACCCTGCTTCTCAGAGCTCATACCTAGAGTAGTGGTGTAGTAGAATAATGCTGA 403
Db 98725 ACATCTAACATTACTAATCTCTAATCCAGGCTCTCCCTGGAATGTGAGCCACATGATC 98666
Qy 404 GCTGCTTATGTCATTTCCAGTTTTTTAGTAGCCACATT-AAAACAGGTAAAAAGGCTGG 462
Db 98665 ACTACATAGGCAATTTAAATTTATCTGTAGGCACATTAAAAATAATAGTAGGCTGG 98606
Qy 463 GCGCAGTGGCTCACCTGTAATCCAGCATTCTGGAGGCTGAGGAGGAGGAGGACATCACCT 522
Db 98605 GTGCAGTGGCTCACACATGTAATCCAGCATTCTCAGAAGCTAAGGTGGGAGATCACTT 98546
Qy 523 TTGCTCAGGAGTTGAGCATTAGCCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 582
Db 98545 GAGGTGAGGAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98487
Qy 583 TACAAAAATTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
Db 98486 CACAAAAATTAGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 98427
Qy 643 GACACAAAGTCACTTAAACCCAGAGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGG 702
Db 98426 GGCAGAGGAATCACTTAAACCTGGGAGACAGAGGTGGCAATAGCAAGATTGAGCCACT 98367
Qy 703 -CACTCCAACTGGGAGACAGAGTGACACTTTTGTCTCRAAAGAAAAAAGAAAAAAGT 761
Db 98366 GAATCCCGGCTGGGTGACAGAGCAAACTCCGCTCTCAAAAAAAGTAAAGTAATTTT 98307
Qy 762 AAAAAGAAACAGGTGAAGTTAACTTTAATAACCAATGTATCCCAATAACAATCAATTC 821
Db 98306 AACCA-----TATATTTATGGACTCAATATATCAATATATATCAATTC 98261
Qy 822 AAAGTGAATTAATATAAACAATTAATGAATGAGATCTTTTACATCTTTTCTTTCTTTTCT 881
Db 98260 AATATGTAATCAATATATAAAATGTAATTAATGAGGTGTTTGCACACTCTT-----TTTTC 98207
Qy 882 ATATTAGTCTTTGAAAGTGAGTATATATGTTTGTGACAGCACATCTCAATTTGGACT 941
Db 98206 ATACTAAGTGTCCTCAAAATCTAATATGCAATTTTCAATTTGAGCAGACATTTCAATTTAGACT 98147
Qy 942 AGCTACATTTTCCAGGTGCTCAGTAGCCACATGTGCTAGCAGTTACTGTATTGGATGGCAC 1001
Db 98146 AGCTACATTTTACCTGCTCAGAGCCACATGCGGCTGGTGGCTAGGTACTGTCCAGCAC 98087

RESULT 15
AC025163
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-533F2, WORKING DRAFT SEQUENCE,
17 unordered pieces.
ACCESSION
AC025163
VERSION
AC025163.17 GI:14547360
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 177876)

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Albrooks, S.L., Amaraturunge, H.C., Are, J.R., Banks, T., Barbacia, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, K., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsged, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwono, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scher, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telirod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

TITLE

Direct Submission

Unpublished

2 (bases 1 to 177876)

Worley, K.C.

REFERENCE

Direct Submission

Submitted (07-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 25, 2001 this sequence version replaced gi:12831249.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAMH
Center clone name: RP11-533F2
----- Sequencing Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 1% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 169802 bases at least Q40
Consensus quality: 180431 bases at least Q30
Consensus quality: 185136 bases at least Q20
Estimated insert size: 177736; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 4.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 42878: contig of 42878 bp in length
* 42879 42978: gap of unknown length
* 42979 62018: contig of 19040 bp in length
* 62019 62118: gap of unknown length
* 62119 80603: contig of 18485 bp in length
* 80604 80703: gap of unknown length
* 80704 80788: contig of 17585 bp in length
* 80789 98388: gap of unknown length
* 98389 113934: contig of 15546 bp in length
* 113935 114034: gap of unknown length
* 114035 125339: contig of 11305 bp in length
* 125340 125439: gap of unknown length
* 125440 138329: contig of 12890 bp in length
* 138330 138429: gap of unknown length
* 138430 145577: contig of 7148 bp in length
* 145578 145677: gap of unknown length
* 145678 150952: contig of 5275 bp in length
* 150953 151052: gap of unknown length
* 151053 155923: contig of 4871 bp in length
* 155924 156024: gap of unknown length
* 156024 159749: contig of 3725 bp in length
* 159749 159849: gap of unknown length
* 159849 163147: contig of 3298 bp in length
* 163147 163247: gap of unknown length
* 163247 166718: contig of 3372 bp in length
* 166718 166719: gap of unknown length
* 166719 170864: contig of 4146 bp in length
* 170865 170964: gap of unknown length
* 170965 173099: contig of 2135 bp in length
* 173100 173199: gap of unknown length
* 173200 175517: contig of 2318 bp in length
* 175518 175617: gap of unknown length
* 175618 177876: contig of 2259 bp in length.

FEATURES

Location/Qualifiers

1..177876
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-533F2"

ORIGIN

Query Match 26.0%; Score 260; DB 2; Length 177876;
Best Local Similarity 67.7%; Pred. No. 3.7e-51;
Matches 447; Conservative 0; Mismatches 190; Indels 23; Gaps 5;
Qy 344 ACCCTGTTCTCAGAGCTCATACCTAGAGTAGTGGTGTAGTAGAATAATGCTGA 403
Db ACATCTAACATTACTAATCTAATCCAGGCTGTCCTCGAATGTGAGCACATGATC 22548
Qy 404 GCTGCTTATGTCATTTCCAGTTTTTTAGTAGCCACATT-AAAACAGGTAAAGGCTGG 462
Db ACTACATAGGCAATTTTAAATTTATCTTGTAGGCACATTAAATAATAGTAGGCTGG 22608
Qy 463 CGCGAGTGGCTCACACCTGTAATCCAGGAGCTTTGGGAGGCTGAGGAGGAGATCACCT 522
Db GTGCAGTGGCTCACACATGTAATCCAGGACATTTTCAAGAGCTTAAGGTTGGGAGATCACTT 22668
Qy 523 TTGCTCAGGAGTTTGAGACATAGCTGCTGCCAACATGCGGAACCTGCTCTTAAAAA 582
Db GAGTGTAGGAGTTGAGAGCAGCTGCTGCAACATGGTGAACCCCGTCTCT-CTAAAAA 22727
Qy 583 TACAAAAATTAGCTGCTGCTGCGGGGCGCTGTGTAATCTCAGCTGCTCAGGAGGCGGA 642
Db CACAAAAATTAGCGGGTGTGGTGGCACACGCTGTAGTCTCAGCTACTCAGGAGGCTGA 22787
Qy 643 GACAAAGATCACTTTAAACCCAGGAGGTTCGAGTTCAGTGTGAGTTCGATCGGCCACT 702

```
Db 22788 GGCAGAGGAATCACTTAAACCTGGGAGACAGAGGTTGCAATAAGCCAAAGATTGAGCCACT 22847
Qy 703 -CACTCCAAACCTGGGAGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGT 761
Db 22848 GAACTCCGGCTGGGTGACAGACAAACTCCGCTCTCAAAAAGTAAAGTAATTT 22907
Qy 762 AAAAAAGAAACAGGTGAAGTTAACTTTAATAACCCAATGTATCCCAAATACAATCATTTTC 821
Db 22908 AACGA-----TATATTTTATTGGACTCAATATATCCAAATATTATCATTTTC 22953
Qy 822 AAAGTGTAAATTAATAATAAACAATATGAATGAGATACTTTACATTTCTTTCTTGTTTTC 881
Db 22954 AATATGTAATCAATATATAAAATGTAATAATGAGGTGTTGACACTCTT-----TTTTC 23007
Qy 882 ATATTAAGTCTTTGAAAAGTGAGTATATATGTTATGCTGACAGACACATCTCAATTTGGACT 941
Db 23008 ATACTAAGTGTCCCAATCTTAATATGCAATTTACATTTGACGACACATCTCAATTTAGACT 23067
Qy 942 AGCTACATTTTCAGGTGCTCAGTAGCCACATGTGGCTAGCAGTTTACTGTATTGGATGGCAC 1001
Db 23068 AGCTACATTTTACCTGCTCAGGAGCCACATGCGGCTGGTGGCTACGGTACTGTGCCAGCAC 23127
```

Search completed: July 4, 2005, 09:55:30
Job time : 2892.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 352.271 Seconds

(without alignments)
16821.327 Million cell updates/sec

Title: US-09-936-271c-13_COPY_6000_7000

Perfect score: 1001

Sequence: 1 gacaacaaagagccccaag.....gttactgtattggatggcac 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	3	AA95905
2	1001	100.0	11570	12	ADK52482
3	1001	100.0	11570	13	ADR72623
4	1001	100.0	11570	13	ADR72875
5	252.6	25.2	52216	6	AAH28355
6	252.6	25.2	52216	6	ABL50307
7	248.8	24.9	56098	11	ACN44922
8	240.2	24.0	55827	8	ACA60949
9	240.2	24.0	55827	10	ABX13671
10	240.2	24.0	58337	13	ADS36454
11	240.2	24.0	64423	13	ADS36462
12	235.6	23.5	27189	11	ACN45066
13	227.4	22.7	2953	4	AAK68359
14	227.4	22.7	2953	4	AAK68360
15	227.4	22.7	3377	12	ADQ64563
16	227.4	22.7	13744	4	AAK82406
17	227.4	22.7	13744	4	AAK68361
18	227.4	22.7	23071	4	AA826699
19	227.4	22.7	23071	8	AA98598
20	227.4	22.7	23071	8	ABX74048

21	227.4	22.7	23071	8	ADA44323	Ada44323 Human sec
22	227.4	22.7	23071	10	ADC20712	Adc20712 Human sec
23	227.4	22.7	23071	10	ADF10881	Adf10881 Human sec
24	226.8	22.7	12275	4	AAK85734	Aak85734 Human imm
25	226.2	22.6	133955	11	ACN45170	Acn45170 Human gen
26	225.4	22.5	325791	4	AA843104	Aas43104 Human Oes
27	225.2	22.5	35641	6	ABL64428	Ab164428 Stomach c
28	225.2	22.5	35641	6	ABN95727	Abn95727 Gene #222
29	224.8	22.5	629	12	ADM99987	Adm99987 Human STA
30	224.8	22.5	629	12	ADM99988	Adm99988 Human STA
31	224.6	22.4	110000	11	ACN43998_3	Continuation (4 of
32	223.8	22.4	15765	5	ABA18276	Abal8276 Human ner
33	223.8	22.4	15765	5	ABA19009	Abal19009 Human ner
34	223.8	22.4	15765	5	ABA17911	Abal7911 Human ner
35	223.8	22.4	15765	5	ABA18250	Abal8250 Human ner
36	223.6	22.3	12267	4	AAK85733	Aak85733 Human imm
37	222	22.2	72409	13	ABD33061	Abd33061 Human can
38	221.6	22.1	2591	4	AAH16224	Aah16224 Human cDN
39	221.6	22.1	2591	12	ADL22885	Adl22885 Human MP2
40	221.6	22.1	2591	13	ADR14082	Adr14082 Human NF-
41	220.8	22.1	96587	9	ADA02984	Ada02984 Human MAP
42	220.8	22.1	96587	10	ADB72722	Adb72722 Human MAP
43	220.8	22.1	96587	10	ADC85464	Adc85464 Human Map
44	220.8	22.1	96587	12	ADM74579	Adm74579 Human car
45	220.6	22.0	203264	11	ACN44506	Acn44506 Human gen

ALIGNMENTS

RESULT 1

AAA95905
ID AAA95905 standard; DNA; 11570 BP.

AC AAA95905;

DT 02-FEB-2001 (first entry)

DE Human KLK-L2 gene.

Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;

kallikrein-like protein; serine protease; cytostatic; cancer;

prostrate cancer; ds.
Homo sapiens.

WO200053776-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-CA000258.

PR 11-MAR-1999; 99US-0124260P.

PR 01-APR-1999; 99US-0127386P.

PR 21-JUL-1999; 99US-0144919P.

PA (MOUN) MOUNT SINAI HOSPITAL.

PI Yousef GM, Diamandis EP;

XX WPI; 2000-587440/55.

XX P-ESDB; AAB21296.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L

protein mediated disorders, especially cancer.

PS Claim 1; Page 143-149; 184pp; English.

XX The present sequence is the coding sequence of the human KLK-L2 gene,
which encodes a kallikrein-like protein. Kallikreins and kallikrein-like
proteins are a subgroup of the serine protease enzyme family. They
catalyse the selective cleavage of specific polypeptide precursors to
release peptides with potent biological activity. Nucleic acids encoding

CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the
CC invention which encodes a secreted serine protease and is located at
CC chromosome 19q13.4.

XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 1e-205;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GACACCAAGAGCCCGGAGGTGAGTCCAGGTTCTTCTGTATACCGACCCACTCTGCG	60
Db	6000	GACACCAAGAGCCCGGAGGTGAGTCCAGGTTCTTCTGTATACCGACCCACTCTGCG	6059
Qy	61	CGCCTTCATCTTCTCCACTTCTCATTTGTTCTGTTGACAGTGCACTTCCCTAAGG	120
Db	6060	CGCCTTCATCTTCTCCACTTCTCATTTGTTCTGTTGACAGTGCACTTCCCTAAGG	6119
Qy	121	TCCTCCAGTGCTTGAATATACAGCTGCTTAAGTCAGAAAAGGTGCGAGGATGCTTTACCCGA	180
Db	6120	TCCTCCAGTGCTTGAATATACAGCTGCTTAAGTCAGAAAAGGTGCGAGGATGCTTTACCCGA	6179
Qy	181	GACAGATAGATGACACCAATGTTCTGCGCGGTGACAAAAGCAGGTAGAGACTCTCTGCCAGG	240
Db	6180	GACAGATAGATGACACCAATGTTCTGCGCGGTGACAAAAGCAGGTAGAGACTCTCTGCCAGG	6239
Qy	241	TGAGGACACCTCTCTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC	300
Db	6240	TGAGGACACCTCTCTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC	6299
Qy	301	GTTCGCCAAATTCAGAAATCCAGAAATGCGCAAGACAGTCCAGGACCCCTGTTCTCACAGA	360
Db	6300	GTTCGCCAAATTCAGAAATCCAGAAATGCGCAAGACAGTCCAGGACCCCTGTTCTCACAGA	6359
Qy	361	GCTCATACCTTAGAGTAGTGTTTGTAGTAAATAATGCTGAGCTGCTTAATGCAATTC	420
Db	6360	GCTCATACCTTAGAGTAGTGTTTGTAGTAAATAATGCTGAGCTGCTTAATGCAATTC	6419
Qy	421	CAGTTTTTTAGTACCCACATTAACAGGTAAAGAGCTGGCGCAGTGCTCACACCT	480
Db	6420	CAGTTTTTTAGTACCCACATTAACAGGTAAAGAGCTGGCGCAGTGCTCACACCT	6479
Qy	481	GTAATCCAGCACATTTGGGAGGCTGAGCAGCGCAGATCACCTTTGGTCAGGAGTTTGAGA	540
Db	6480	GTAATCCAGCACATTTGGGAGGCTGAGCAGCGCAGATCACCTTTGGTCAGGAGTTTGAGA	6539
Qy	541	CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAAAATAAAAAATTAGCCTGGC	600
Db	6540	CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAAAATAAAAAATTAGCCTGGC	6599
Qy	601	ATGTTGGCGGGCGCTGTAAATCTCAGCTGCTCAGAGGCCGAGACACAGAAATCACTTAA	660
Db	6600	ATGTTGGCGGGCGCTGTAAATCTCAGCTGCTCAGAGGCCGAGACACAGAAATCACTTAA	6659
Qy	661	ACCAGGAGGTGGAGTTGGAGTGAGTGGTGCACCTCCACTCCAACTGGGAGAC	720
Db	6660	ACCAGGAGGTGGAGTTGGAGTGAGTGGTGCACCTCCACTCCAACTGGGAGAC	6719
Qy	721	AGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGGAAG	780
Db	6720	AGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGGAAG	6779
Qy	781	TTAATTTAATAACCAATGTATCCCAATACATTTCAAGTGTAATTAATAAA	840
Db	6780	TTAATTTAATAACCAATGTATCCCAATACATTTCAAGTGTAATTAATAAA	6839
Qy	841	ACAATATGATGAGATCTTACATCTTCTTCTTTTTCATATTAAGTCTTTGAAAGT	900
Db	6840	ACAATATGATGAGATCTTACATCTTCTTCTTTTTCATATTAAGTCTTTGAAAGT	6899
Qy	901	GAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGCTC	960
Db	6900	GAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGCTC	6959

Qy 961 AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC 1001
|||||
Db 6960 AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC 7000
|||||

RESULT 4

ADR72875
ID ADR72875 standard; DNA; 11570 BP.

XX ADR72875;

XX XX 02-DEC-2004 (first entry)

XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.

XX kallikrein 5; tumour marker; ovarian cancer;

KW epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;

XX KLK5; ds; gene.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 2221..11247

FT FT /*tag= b

FT FT /product= "Human ovarian cancer-related tumour marker

FT FT kallikrein 5 (hk5) protein"

FT FT 2221..2293

FT FT /*tag= a

FT FT /number= 1

FT FT /*tag= c

FT FT /number= 1

FT FT 4762..5023

FT FT /*tag= d

FT FT /number= 2

FT FT 5024..5762

FT FT /*tag= e

FT FT /number= 2

FT FT 5763..6019

FT FT /*tag= f

FT FT /number= 3

FT FT 6020..6104

FT FT /*tag= g

FT FT /number= 3

FT FT 6105..6238

FT FT /*tag= h

FT FT /number= 4

FT FT 6239..11091

FT FT /*tag= i

FT FT /number= 4

FT FT 11092..11247

FT FT /*tag= j

FT FT /number= 5

PN WO2004075713-A2.

XX 10-SEP-2004.

XX 26-FEB-2004; 2004WO-CA000281.

XX 26-FEB-2003; 2003US-0450406P.

XX (MOUN) MOUNT SINAI HOSPITAL.

XX Diamandis EP;

XX WPI; 2004-661815/64.

XX P-ESDB; ADR72875.

XX Kallikrein markers detection method for detecting ovarian cancer in
PT patient, involves detecting kallikrein markers and CA125 in sample
PT obtained from patient, and comparing detected amounts with standard

amounts.

Example 2; SEQ ID NO 5; 102pp; English.

The invention relates to a novel method for detecting a plurality of kallikrein markers associated with ovarian cancer. The method comprises obtaining a sample from a patient and detecting in the sample a plurality of kallikrein markers, and optionally carbohydrate antigen CA125, wherein the kallikrein markers are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11. The detected amounts of the kallikrein markers are compared with standard amounts. The method of the invention may be useful for detecting kallikrein markers associated with ovarian cancer in a patient and thus for detecting ovarian cancer, particularly epithelial ovarian carcinoma. The current sequence is that of the human ovarian cancer-related tumour marker kallikrein 5 (KUK5) DNA 2 of the invention which encodes a secreted serine protease and is located at chromosome 19q13.4.

Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1001;	DB 13;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 1e-205;		
Matches 1001;	Conservative -0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GACAA	CCAGAGCC	CCCAAGGTGAGTGTCCAGGTTCTTCTTGATACGAGCCCATCTCTGC 60
Db	6000	GACAAC	CAAGAGCC	CCCAAGGTGAGTGTCTTCTTGATACGAGCCCATCTCTGC 6059
Qy	61	CGCCTT	CCATCTTT	CTCCACTTCTCAATGTGTCTCTTTTGACAGTGCATCTCCCTAAGG 120
Db	6060	CGCCTT	CCATCTTT	CTCCACTTCTCAATGTGTCTCTTTTGACAGTGCATCTCCCTAAGG 6119
Qy	121	TCCTC	GAGTGC	TGGAATATCAGCGTGTAAGTCAGAAAGGTGGAGGATGCTTACC CGA 180
Db	6120	TCCTC	GAGTGC	TGGAATATCAGCGTGTAAGTCAGAAAGGTGGAGGATGCTTACC CGA 6179
Qy	181	GACAG	ATAGATG	ACACCATGTTCTGCGCCGCTGACAAAGCAGGTAGAGATCTCTGCCAGG 240
Db	6180	GACAG	ATAGATG	ACACCATGTTCTGCGCCGCTGACAAAGCAGGTAGAGATCTCTGCCAGG 6239
Qy	241	TGAGG	ACACCTCT	CTTTTATTCAGCAGATACACACTGAGTGCCCACTCGGTAAACATGGAGC 300
Db	6240	TGAGG	ACACCTCT	CTTTTATTCAGCAGATACACACTGAGTGCCCACTCGGTAAACATGGAGC 6299
Qy	301	GTGGC	CAAAAT	CTGAGAAATCCAGCAATGCGACAGACAGTCAGGACCCCTGTTCTCAGAG 360
Db	6300	GTGGC	CAAAAT	CTGAGAAATCCAGCAATGCGACAGACAGTCAGGACCCCTGTTCTCAGAG 6359
Qy	361	GCTCA	TACCCCT	TAGAGTAGTGGTGTATAGTAGAATAATGCTGAGCTGCTTATGTCAATTC 420
Db	6360	GCTCA	TACCCCT	TAGAGTAGTGGTGTATAGTAGAATAATGCTGAGCTGCTTATGTCAATTC 6419
Qy	421	CAGT	TTTTTT	TAGTAGCCACATTTAAACAGGTAAAAAAGCCTGGCGCAGTGGCTCACACCT 480
Db	6420	CAGT	TTTTTT	TAGTAGCCACATTTAAACAGGTAAAAAAGCCTGGCGCAGTGGCTCACACCT 6479
Qy	481	GTAAT	CCAGC	ACCTTTGGGAGGCTGAGGCGGAGAGATCACTTTTGGTGCAGAGTTTGAGA 540
Db	6480	GTAAT	CCAGC	ACCTTTGGGAGGCTGAGGCGGAGAGATCACTTTTGGTGCAGAGTTTGAGA 6539
Qy	541	CTAGC	CTGGCC	CAACATCGCGAAACTCTGTCTCTAAAAAAAATAACAAAAATTAGCCCTGGC 600
Db	6540	CTAGC	CTGGCC	CAACATCGCGAAACTCTGTCTCTAAAAAAAATAACAAAAATTAGCCCTGGC 6599
Qy	601	ATG	TGTGCG	GGCGGCGCTGTAAATCTCAGCTGTCTCAGGAGCGCCGAGACACAAGAATCACCTTAA 660
Db	6600	ATG	TGTGCG	GGCGGCGCTGTAAATCTCAGCTGTCTCAGGAGCGCCGAGACACAAGAATCACCTTAA 6659
Qy	661	ACC	CAGAG	GTGGAGTTGAGTGAGCTGAGATCGTGCACTCACTCAACCTGGGAGAC 720
Db	6660	ACC	CAGAG	GTGGAGTTGAGTGAGCTGAGATCGTGCACTCACTCAACCTGGGAGAC 6719

Qy	721	AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAAGAAACACAGGTGAAG	780
Db	6720	AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAAGAAACACAGGTGAAG	6779
Qy	781	TTAACTTTTAATAACCCCAATGTATCCCAATACAAATCATTTTCAAAGTGTAAATTAATATAA	840
Db	6780	TTAACTTTTAATAACCCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATTAATATAA	6839
Qy	841	ACAATTATGAATGAGATACCTTTACATCTCTTTCTTGTCTTTTCATATTAAGTCTTTGAAAGT	900
Db	6840	ACAATTATGAATGAGATACCTTTACATCTCTTTCTTGTCTTTTCATATTAAGTCTTTGAAAGT	6899
Qy	901	GAGTATATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTCAAGTGTCTC	960
Db	6900	GAGTATATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTCAAGTGTCTC	6959
Qy	961	AGTAGCCACATGCGGTAGCAGTTACTGTATTGGATGGCAC	1001
Db	6960	AGTAGCCACATGCGGTAGCAGTTACTGTATTGGATGGCAC	7000
RESULT 5			
AAH28355			
ID	AAH28355 standard; DNA; 52216 BP.		
XX	AAH28355;		
XX	05-SEP-2001 (first entry)		
XX	Nucleotide sequence of the human musashi promoter.		
DE			
XX	Musashi promoter; multipotential neural progenitor cell;		
KW	neural stem cell; central nervous system; ss.		
KW			
XX	Homo sapiens.		
OS			
XX	WO200146384-A2.		
FN			
XX	28-JUN-2001.		
PD			
XX	22-DEC-2000; 2000WO-US035395.		
PF			
XX	23-DEC-1999; 99US-0173003P.		
PR	(CORR) CORNELL RES FOUND INC.		
XX	(NISC-) JAPAN SCI & TECHNOLOGY CORP.		
PA			
PA	Goldman SA, Okano H;		
XX	WPI; 2001-418053/44.		
FI			
XX	Separating multipotential neural progenitor cells from a mixed population		
XX	of cells, involves introducing nucleic acid molecule encoding fluorescent		
PT	protein under promoter control, and separating fluorescent cells.		
PT			
XX	Claim 12; Fig 11A-JJ; 87pp; English.		
PS			
XX	The present sequence represents the human musashi promoter, which is used		
CC	in the method of the invention. The specification describes a method for		
CC	separating multipotential neural progenitor cells from a mixed population		
CC	of cell types. The method comprises introducing a nucleic acid encoding a		
CC	fluorescent protein under control of a promoter which selectively		
CC	functions in the progenitor cells, into all cell types of the mixed		
CC	population, allowing only the progenitor cells to express the fluorescent		
CC	protein, and identifying and separating cells that are fluorescent, where		
CC	the separated cells are progenitor cells. The method is useful for		
CC	isolation and purification of multipotential neural progenitor cells,		
CC	especially neural stem cells from adult brain. The isolated cells are		
CC	used in both basic analyses of precursor and stem cell growth control, as		
CC	well as in more applied studies of their transplantability and		
CC	engraftment characteristics. The cells are useful in support of the		
CC	structural repair of the damaged central nervous system, such as in the		
CC	traumatized brain, or the contoured,traumatized or transected spinal		

The present sequence represents the human musashi promoter, which is used in the method of the invention. The specification describes a method for separating multipotential neural progenitor cells from a mixed population of cell types. The method comprises introducing a nucleic acid encoding a fluorescent protein under control of a promoter which selectively functions in the progenitor cells, into all cell types of the mixed population, allowing only the progenitor cells to express the fluorescent protein, and identifying and separating cells that are fluorescent, where the separated cells are progenitor cells. The method is useful for isolation and purification of multipotential neural progenitor cells, especially neural stem cells from adult brain. The isolated cells are used in both basic analyses of precursor and stem cell growth control, as well as in more applied studies of their transplantability and engraftment characteristics. The cells are useful in support of the structural repair of the damaged central nervous system, such as in the traumatized brain, or the contoured, traumatized or transected spinal

[illegible]

```
RESULT 7
ACN44922/c
ID ACN44922 standard; DNA; 56098 BP.
XX AC
XX ACN44922;
XX
XX 18-NOV-2004 (first entry)
DT
XX Human genomic sequence HCG1779020.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX KW
XX Homo sapiens.
XX OS
XX WO2003073826-A2.
XX FN
XX PD
XX 12-SEP-2003.
XX PF
XX 28-FEB-2003; 2003WO-US006235.
XX PR
XX 01-MAR-2002; 2002US-00087192.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI
XX Morris DW;
XX
XX WPI; 2003-328604/31.
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 1612; opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published
XX
XX Sequence 56098 BP; 17842 A; 10579 C; 10322 G; 16942 T; 0 U; 413 Other;
XX
XX Query Match 24.9%; Score 248.8; DB 11; Length 56098;
XX Best Local Similarity 71.0%; Pred. No. 1.2e-43;
XX Matches 402; Conservative 0; Mismatches 152; Indels 12; Gaps 5;
XX
XX 440 TTAACACAGGTAAAGGCTGGCGCAGTGGCTCACACCTGTATATCCAGCAGCTTTGGG 499
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 26680 TTAAAAAAGTATGATGACAGCCAGGCGTGTGGCTCACACCTGTATATCCAGCAGTTGGG 26621
XX
XX 500 AGGCTGAGCAGGAGATACCTTTGGTCAGGAGTTTGAGACTAGCTGGCCACATGCC 559
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 26620 AGGCTGAGTGGGCGAGATCACTTTGAGTGGTACGGAATTTGAAACCGACCTGGCCACATGGT 26561
XX
XX 560 GAAACTCTGTCTCTAAAAAATAACAAAAATAGCCTGGCATGTGGCGGCGCTCTGA 619
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 26560 GAAACCCCATCTCT-ACTAAATAATAGAAAGTTAGCCGGATGTGGTGACACAGCTTGTGA 26502
XX
XX 620 ATCTCAGCTGTCTCAGGAGCGCCAGACACAAAGATCACTTAAACCCAGAGGTGGAGTTG 679
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 26501 GTCCAGACGCTCGGAGGAGGAGGACAAAGATCACTTGAACCGCAAGTGGAGGTTG 26442
XX
XX 680 CAGTGAGCTGAGATCGTGCCA-CTCACTCCAACTGGGAGACAGAGTGCACATTTTGCT 738
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 26441 CAGTGAGCGGAGATTGAGGTACCCACTCCAGCCTGGCGACAGAGGAGACTCACTC 26382
Qy 739 CAAAAGAGAAAAAAGAAAGTAAAGAAAGAAAGAGAGTGAAGTTAACTTTTAATACCCAA 798
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26381 AGAAACAAAAATAAAA-----CAAAACAAAAAAGTGAATAATGTTAAATTTTAAACAATAAA 26328
Qy 799 TGATATCCCAATACAAATCAATTTCAAGTGTAAATTAATATAAACAATATGAATGAGATA 858
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26327 TTTATTTAAACCGTTATATATCTAAATGTCAATTTGAACATGTAATCAATGTGAACCTTTTA 26268
Qy 859 CTTTACATCTTTTCTTTGTTTTCATA-TTAAGTCTTTTGAAGTGAAGTATATATGTTATGC 917
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26267 ATGGAATATTTACATCTTTTAAAGATAAGTCTTTGAAGTCCAGTGTGTTATATACAC 26208
Qy 918 TGACAGACACA---TCTCAATTTGGACTAGCTACATTTTCAGGTGCTCAGTAGCCACATGTG 974
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26207 TTATAGCACAACTTCTCAATTTGGACTAGCCACATTTCAAGTGTCTCAACAGGCACATGTG 26148
Qy 975 GCTAGCAGTTACTGTATTGGATGGCA 1000
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26147 GCCAGTGGCTACTGTTTGGGATAACA 26122

RESULT 8
ACA60949/c
ID ACA60949 standard; DNA; 55827 BP.
XX AC
XX ACA60949;
XX
XX 11-AUG-2003 (first entry)
DT
XX DNA encoding human carboxypeptidase.
DE
XX Human; gene; carboxypeptidase; inflammation; cancer; arteriosclerosis;
XX neurodegenerative disease; protease; ds; single nucleotide polymorphism;
XX SNP.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX variation replace(858,T)
XX /*tag= u
XX /note= "Single nucleotide polymorphism"
XX CDS 2017..53409
XX /*tag= a
XX /product= "Carboxypeptidase"
XX exon 2017..2090
XX /*tag= b
XX /number= 1
XX intron 2091..19206
XX /*tag= c
XX
XX variation replace(2122,T)
XX /*tag= v
XX /note= "Single nucleotide polymorphism"
XX variation replace(4088,A)
XX /*tag= w
XX /note= "Single nucleotide polymorphism"
XX variation replace(4240,A)
XX /*tag= x
XX /note= "Single nucleotide polymorphism"
XX variation replace(4416..4418,TA)
XX /*tag= y
XX /note= "Single nucleotide polymorphism"
XX variation replace(4434,G)
XX /*tag= z
XX /note= "Single nucleotide polymorphism"
XX variation replace(4969,C)
XX /*tag= aa
XX /note= "Single nucleotide polymorphism"
XX variation replace(5868,A)
XX /*tag= ab
XX /note= "Single nucleotide polymorphism"
XX FT
```

FT	variation	replace(6254,C)	FT	variation	/*tag= as
FT		/*tag= ac	FT		/note= "Single nucleotide polymorphism"
FT	variation	replace(10171,C)	FT	variation	replace(31600,C)
FT		/*tag= ad	FT		/*tag= at
FT	variation	/*tag= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(10452,A)	FT		replace(31714,C)
FT		/*tag= ae	FT		/*tag= au
FT	variation	/note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(11613,A)	FT		replace(31715,T)
FT		/*tag= af	FT		/*tag= av
FT	variation	/*tag= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(12130,T)	FT		replace(32193,G)
FT		/*tag= ag	FT		/*tag= aw
FT	variation	/note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(17867. .17869,GA)	FT		replace(32341,A)
FT		/*tag= ah	FT		/*tag= ay
FT	variation	/note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(18243,G)	FT		replace(32341,G)
FT		/*tag= ai	FT		/*tag= ax
FT	variation	/*tag= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(19207. .19282	FT		replace(32561,G)
FT	exon	/*tag= d	FT		/*tag= az
FT		/number= 2	FT		/note= "Single nucleotide polymorphism"
FT	intron	19283. .22683	FT	variation	replace(32600. .32602,AA)
FT		/*tag= e	FT		/*tag= ba
FT		/number= 2	FT	variation	/note= "Single nucleotide polymorphism"
FT	variation	replace(22450,C)	FT		replace(32642,T)
FT		/*tag= aj	FT		/*tag= bb
FT		/note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT	exon	22684. .22808	FT		replace(32793. .32795,TG)
FT		/*tag= f	FT		/*tag= bc
FT		/number= 3	FT	exon	/note= "Single nucleotide polymorphism"
FT	intron	22809. .24477	FT		33053. .33157
FT		/*tag= g	FT		/*tag= l
FT		/number= 3	FT	variation	/number= 6
FT	variation	replace(23003,T)	FT		replace(33071,A)
FT		/*tag= ak	FT		/*tag= bd
FT	variation	/note= "Single nucleotide polymorphism"	FT	intron	/note= "Single nucleotide polymorphism"
FT		replace(24055. .24057,GA)	FT		33158. .42288
FT		/*tag= al	FT		/*tag= m
FT		replace(24132,T)	FT	variation	/number= 6
FT	variation	/note= "Single nucleotide polymorphism"	FT		replace(34721,T)
FT		/*tag= am	FT		/*tag= be
FT	exon	24478. .24586	FT	variation	/note= "Single nucleotide polymorphism"
FT		/*tag= h	FT		replace(35304,A)
FT		/number= 4	FT	variation	/*tag= bf
FT	intron	24587. .28129	FT		/note= "Single nucleotide polymorphism"
FT		/*tag= i	FT		replace(35425,C)
FT		/number= 4	FT	variation	/*tag= bg
FT	variation	replace(25713,A)	FT		/note= "Single nucleotide polymorphism"
FT		/*tag= an	FT	variation	/*tag= bh
FT		/note= "Single nucleotide polymorphism"	FT		/note= "Single nucleotide polymorphism"
FT	exon	28130. .28231	FT	variation	replace(36291,G)
FT		/*tag= j	FT		/*tag= bi
FT		/number= 5	FT	exon	/note= "Single nucleotide polymorphism"
FT	intron	28232. .33052	FT		42289. .42382
FT		/*tag= k	FT		/*tag= n
FT		/number= 5	FT	intron	/number= 7
FT	variation	replace(28476,C)	FT		42383. .48643
FT		/*tag= ao	FT		/*tag= o
FT	variation	/note= "Single nucleotide polymorphism"	FT	exon	/number= 7
FT		replace(29404,G)	FT		48644. .48846
FT		/*tag= ap	FT		/*tag= p
FT	variation	/note= "Single nucleotide polymorphism"	FT		/number= 8
FT		replace(31047,T)	FT	intron	48847. .51176
FT		/*tag= aq	FT		/*tag= q
FT		/note= "Single nucleotide polymorphism"	FT	exon	/number= 8
FT	variation	replace(31445,C)	FT		51177. .51264
FT		/*tag= ar	FT		/*tag= r
FT		/note= "Single nucleotide polymorphism"	FT	intron	/number= 9
FT	variation	replace(31447,C)	FT		51265. .53224
FT			FT		/*tag= s


```
FT FT /cons_splice= (5'site:yes,3'site:no)
FT FT replace(23003,T)
FT FT /*tag= ak
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(24055..24056,GAA)
FT FT /*tag= al
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(24132,C)
FT FT /*tag= am
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 24478..24586
FT FT /*tag= h
FT FT /number= 4
FT FT 24587..28129
FT FT /*tag= i
FT FT /number= 4
FT FT /cons_splice= (5'site:no,3'site:no)
FT FT replace(25713,A)
FT FT /*tag= an
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 28130..28231
FT FT /*tag= j
FT FT /number= 5
FT FT 28232..33052
FT FT /*tag= k
FT FT /number= 5
FT FT replace(28476,C)
FT FT /*tag= ao
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(29404,G)
FT FT /*tag= ap
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31047,T)
FT FT /*tag= aq
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31445,C)
FT FT /*tag= ar
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31447,C)
FT FT /*tag= as
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31600,C)
FT FT /*tag= at
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31714,C)
FT FT /*tag= au
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31715,T)
FT FT /*tag= av
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32193,G)
FT FT /*tag= aw
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32341,R)
FT FT /*tag= ax
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32561,G)
FT FT /*tag= ay
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32600..35601,AAA)
FT FT /*tag= az
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32642,T)
FT FT /*tag= ba
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32793..32794,TAG)
FT FT /*tag= bb
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 33053..33157
FT FT /*tag= l
FT FT /number= 6
FT FT replace(33071,A)
FT FT /*tag= bc

FT intron /standard_name= "Single nucleotide polymorphism"
FT FT 33158..42288
FT FT /*tag= m
FT FT /number= 6
FT FT /cons_splice= (5'site:no,3'site:no)
FT FT replace(34721,T)
FT FT /*tag= bd
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(35304,A)
FT FT /*tag= be
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(35425,C)
FT FT /*tag= bf
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(36050,G)
FT FT /*tag= bg
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(36291,G)
FT FT /*tag= bh
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 42289..42382
FT FT /*tag= n
FT FT /number= 7
FT FT 42383..48643
FT FT /*tag= o
FT FT /number= 7
FT FT /cons_splice= (5'site:yes,3'site:no)
FT FT 48644..48846
FT FT /*tag= p
FT FT /number= 8
FT FT 48847..51176
FT FT /*tag= q
FT FT /number= 8
FT FT /cons_splice= (5'site:yes,3'site:no)
FT FT 51177..51264
FT FT /*tag= r
FT FT /number= 9
FT FT 51265..53224
FT FT /*tag= s

Query Match 24.0%; Score 240.2; DB 10; Length 55827;
Best Local Similarity 68.6%; Pred. No. 8.6e-42;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

Qy 363 TCATACCTAGAGTAGGTGGTGTAGTAGAATAATGCTGAGCTGCTATGTCATTTCCTCA 422
Db 46943 TGAATCCTCAGACCAGGGTGTGTTAATAGAAATATAATGCAAGTCACATATTTACTTAA 46884

Qy 423 GTTTTGTAGTAGGCACATTAACAGGTAAAA--AAGGCTGGGGCGCAGTGGCTCACACCT 480
Db 46883 ATTTTCAAGTTACCACTTTAAGAAAAATAAAAAAGAGCGGCAGGTCGCTCACACCT 46824

Qy 481 GTAATCCAGCACCTTTGGGAGGCTGAGGAGGAGATCACTTTGGTCAGAGTTTGAGA 540
Db 46823 ATAATCCAGCACCTTTGGGAGGCTGAGGAGGCGGATCCTTGAGATCAGAGTTTGACA 46764

Qy 541 CTAGCCTGGCCACATGCGGAAACTCTGTCTCTAAAAAATAACAAAAATTAGCCTGGC 600
Db 46763 CCAGCAGGCGCACATGCTGAAACCCGCTCTTACCAAAA--TACAAAAATAGCCAGGC 46706

Qy 601 ATGCTGGCGGCGCCTGTAATCTCAGCTGCTCAGAGGCGGAGACACAAGAAATCACTTAA 660
Db 46705 ATAGTGTGCAGCGCTGTAATCCAGCTACTTTGGGAGGCTGAGGACGAAACTTGTCTGA 46646

Qy 661 ACCCAGAGGTGGAGGTTTGCAGTGAGCTGCTGCCACT--CACTCCACCTGGGAGA 719
Db 46645 ACCCAGAGGCGAGGTTTGCAGTGAGCTGAGTATGATCATGACACTGCACCTCCAGCTGGATGA 46586

Qy 720 CAGAGTGACA--CTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA 778
Db 46585 CAGAACAGACCCCTGCCACACACACACACACAAAGAAATATGTGAGATTAATTCATA 46526

Qy 779 AGTTAACTTTTAATAACCAATGTATCCCAATAACAATCAATTCATTAATTAATA 838
```


PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0198874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216664P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229348P.
PR 01-SEP-2000; 2000US-0229349P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR PA (HUMA-) HUMAN GENOME SCI INC.
PR PI Rosen CA, Barash SC, Ruben SM;
PR XX WPI; 2001-483426/52.
PR XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PR PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PR XX Disclosure; SEQ ID NO 23172; 3071pp + Sequence Listing; English.
PR XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 2953 BP; 860 A; 572 C; 619 G; 902 T; 0 U; 0 Other;

Query Match 22.7%; Score 227.4; DB 4; Length 2953;
Best Local Similarity 69.1%; Pred. No. 2.9e-39;
Matches 403; Conservative 0; Mismatches 161; Indels 19; Gaps 6;
Qy 414 TCATTTCCAGTGTCTTTTAGTAGCCACATTAAACACAGGTAATAAAA-----GGCTGGGCGCA 467
Db 1699 TAATTAAGATTTTAAACAGCCACATTATTAATGTAAGAAAAAGTCGGCGCGCG 1640
Qy 468 GTGGCTCACACTGTAAATCCAGACACTTTGGGAGGCTGAGCAGCGAGATCACCTTTGGT 527
Db 1639 GTGGCTCACGCTGTAAATCCAGACACTTTGGGAGACCGAGCAGCGGATCAC--GACGT 1582
Qy 528 CAGGAGTTTGAGACTAGCTCGCCCAACATGCGAACTCTCTCTAATAAAAAATACAA 587
Db 1581 CGAAGATCGAGACCATCTTTGTAACACGCGTAACCTCGCTCTACTAATAAATACAA 1522
Qy 588 AAATTAGCTGGCATGTGGGCGCGCTGTAATCTCAGCTGCTCAGAGGCGCGAGAC 647
Db 1521 AAATTAGCGGATGTGGCGGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCG 1462
Qy 648 AAGAACTACTTAAACCCAGAGGTGGAGTTGCGAGTGAGCTGAGATCGTCCACTCACTC 707
Db 1461 GAGAATGGCGTCAACCCGAGGAGAGAGCTTGCAGTGAGCTGAGATGACCACTGCACT 1402
Qy 708 CAACCTGGGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAATAAAA 767
Db 1401 CCACGAGCTGGGCGACAGAGAC--AGACTCTGTCTCAGAAAAAAGAAAAAAGAAAA 1343
Qy 768 GAAACAGGTGAAGTTAACTTTAAATCAACCAATGATCCCAATACAAATCAATTTCAAAGTG 827
Db 1342 TAAGTCATAGATTTTAATATATTAGCTCAGTATGCTTAAATATATATATTTAA----- 1288
Qy 828 TAATTAATATAAACAATATGAATGAGATCTTTTACATTTCTTTTCTTTTTCATATTA 887
Db 1287 -CATGTAATCAACAAAGTTAATAACAAAATATTCTACATTCCTTGTGGGGTAAT--A 1232
Qy 888 AGTCTTTGAAGTGTATATATG--TTATGCTGACAGCAGATCTCAATTTGGACTAGCTA 946
Db 1231 AATCTTTGAAATTAATATGTTATTTTATCTTACAGCTGATCTTGATTCAGACTAGCCA 1172
Qy 947 CATTTTCAGTGTCTAGTAGCCACATGTGGCTAGCAGATTACTGT 989
Db 1171 CATTTTCAGTGTCTAGTAGCCCATGTGGCTAGCACTATGTT 1129

RESULT 15

ADQ64563/c

ID ADQ64563 standard; cDNA; 3377 BP.

XX

AC ADQ64563;

XX

DT 07-OCT-2004 (first entry)

XX

DE Novel human cDNA sequence #1724.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytosstatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.

OS Homo sapiens.

XX EPI440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

PR 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.
DR P-PSDB; ADQ66751.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 1724; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.

XX SQ Sequence 3377 BP; 937 A; 660 C; 718 G; 1062 T; 0 U; 0 Other;

Query Match 22.7%; Score 227.4; DB 12; Length 3377;
Best Local Similarity 69.1%; Pred. No. 2.9e-39;

Matches 403; Conservative 0; Mismatches 161; Indels 19; Gaps 6;

Qy 414 TCATTTCCAGTGTCTTTTAGTAGCCACATTAAACACAGGTAATAAAA-----GGCTGGGCGCA 467
Db 2132 TAATTAAGATTTTAAACAGCCACATTATTAATGTAAGAAAAAGTCGGCGCGCG 2073

Qy 468 GTGGCTCACACTGTAAATCCAGACACTTTGGGAGGCTGAGCAGCGAGATCACCTTTGGT 527
Db 2072 GTGGCTCACGCTGTAAATCCAGACACTTTGGGAGACCGAGCAGCGGATCAC--GACGT 2015

Qy 528 CAGGAGTTTGAGACTAGCTGCGCAACATGCGAACTCTGTCTTAAAAAATAACAA 587
Db 2014 CGGAAGATCGAGACCATCTTTGTTAACCGGTGAACTCCGTCTCTACTAAAAATACAA 1955

Qy 588 AAATTAGCTGGCATGTGGGCGCGCTGTAATCTCAGCTGTCTCAGAGGCGCGAGAC 647
Db 1954 AAATTAGCGGATGTGGCGGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCG 1895

Qy 648 AAGAACTACTTAAACCCAGAGGTGGAGTTGCGAGTGAGATCGTCCACTCACTC 707
Db 1894 GAGAATGGCGTGAACCCGAGGAGCGAGCTTGCAGTGAGCTGAGATGACCACTGCACT 1835

Qy 708 CAACCTGGGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAATAAAA 767
Db 1834 CCACGAGCTGGGCGACAGAGAC--AGACTCTGTCTCAGAAAAAAGAAAAAAGAAAA 1776

Qy 768 GAAACAGGTGAAGTTAACTTTAATTAACCAATGATCCCAATACAAATCAATTTCAAAGTG 827
Db 1171 CATTTTCAGTGTCTAGTAGCCCATGTGGCTAGCACTATGTT 1129

```
Db 1775 TAAGTCATAGATTTTAAATATATTTAGCTCAGTATGCTAAATAATTAATCAATTTAA----- 1721
Qy 828 TAATTAATATATAAAACAATATGAATGAGATCTTTTACATTCCTTTCTGTTTTCATATTA 887
Db 1720 -CATGTAATCAACAAGTTATTACAAATATTTACAAATCTTGTGGGGTAAT---A 1665
Qy 888 AGTCTTTGAAAGTGAGTATATATG-TTATGCTGACAGCACATCTCAATTTGGACTAGCTA 946
Db 1664 AATCTTTGAAATTAATGTAATTTTATACCTACAGCTGATCTTGATTCAGACTAGCCA 1605
Qy 947 CATTTCAGGTGCTCAGTAGCCACATGTGGCTAGCAGTTACTGT 989
Db 1604 CATTTCCAGTGCTCAGTAGCCCCCATGTGGCTAGCAACTATGTT 1562
```

Search completed: July 4, 2005, 03:03:25
Job time : 358.271 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	1001	100.0	10818	4	US-09-949-016-13583	Sequence 13583, A
C 2	600.6	60.0	601	4	US-09-949-016-63399	Sequence 63399, A
C 3	600.6	60.0	601	4	US-09-949-016-63400	Sequence 63400, A
C 4	600.6	60.0	601	4	US-09-949-016-63401	Sequence 63401, A
C 5	546	54.5	601	4	US-09-949-016-63402	Sequence 63402, A
C 6	254	25.4	601	4	US-09-949-016-63403	Sequence 63403, A
C 7	240.2	24.0	55827	3	US-09-813-133A-3	Sequence 3, Appli
C 8	240.2	24.0	55827	4	US-10-212-877-3	Sequence 3, Appli
C 9	235.6	23.5	75674	4	US-09-949-016-17597	Sequence 17597, A
C 10	230.2	23.0	265038	4	US-09-949-016-15779	Sequence 15779, A
C 11	229.8	23.0	131724	4	US-09-949-016-12893	Sequence 12893, A
C 12	226.2	22.6	109159	4	US-09-949-016-14169	Sequence 14169, A
C 13	226.2	22.6	109159	4	US-09-949-016-14170	Sequence 14170, A
C 14	225.6	22.5	18319	4	US-09-949-016-17446	Sequence 17446, A
C 15	225.4	22.5	325791	4	US-09-768-185A-1	Sequence 1, Appli
C 16	225.2	22.5	601	4	US-09-949-016-202742	Sequence 202742, A
C 17	224.6	22.4	72992	4	US-09-949-016-17592	Sequence 17592, A
C 18	221.4	22.1	70131	4	US-09-949-016-16064	Sequence 16064, A
C 19	220.8	22.1	209210	4	US-09-949-016-15094	Sequence 15094, A
C 20	220.2	22.0	150394	4	US-09-949-016-13042	Sequence 13042, A
C 21	220	22.0	136264	4	US-09-949-016-12756	Sequence 12756, A
C 22	220	22.0	136265	4	US-09-949-016-13001	Sequence 13001, A
C 23	219.8	22.0	166698	4	US-09-949-016-16038	Sequence 16038, A
C 24	219.4	21.9	35058	4	US-09-949-016-12607	Sequence 12607, A
C 25	219.4	21.9	35059	4	US-09-949-016-13831	Sequence 13831, A
C 26	219.4	21.9	165651	4	US-09-949-016-13032	Sequence 13032, A
C 27	219.2	21.9	16984	4	US-09-949-016-16893	Sequence 16893, A

Db 3545 GTTGCCAAATCTGAGAATCCAGAAATGCGCAACAGACAGTCCAGACCCCTGCTTCACAGA 3604
Qy 361 GCTCATACCTAGAGTAGTGGTGTAGTAAATATGCTGAGCTGCTTATGTCATTTTC 420
Db 3605 GCTCATACCTAGAGTAGTGGTGTAGTAAATATGCTGAGCTGCTTATGTCATTTTC 3664
Qy 421 CAGTTTTTTAGTAGCCACATTTAAACAGAGTAAAGAGCTGGCGCAGTGGCTCACACCT 480
Db 3665 CAGTTTTTTAGTAGCCACATTTAAACAGAGTAAAGAGCTGGCGCAGTGGCTCACACCT 3724
Qy 481 GTAATCCAGACATTTTGGGAGGCTGAGCGAGCAGATCACTTTTGGTCAGAGTTTGAGA 540
Db 3725 GTAATCCAGACATTTTGGGAGGCTGAGCGAGCAGATCACTTTTGGTCAGAGTTTGAGA 3784
Qy 541 CTAGCCTGGCCACATCGCGAAACTCTGCTCTTAAAGAAATACAAAGAAATAGCCTGGC 600
Db 3785 CTAGCCTGGCCACATCGCGAAACTCTGCTCTTAAAGAAATACAAAGAAATAGCCTGGC 3844
Qy 601 ATGCTGGCGGCGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACACAGAAATCACCTTAA 660
Db 3845 ATGCTGGCGGCGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACACAGAAATCACCTTAA 3904
Qy 661 ACCAGAGGTGGAGGTTCAGTGAGTGCCTGCTCACTCCAACTGGGAGAC 720
Db 3905 ACCAGAGGTGGAGGTTCAGTGAGTGCCTGCTCACTCCAACTGGGAGAC 3964
Qy 721 AGAGTGACATTTTGTCTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 780
Db 3965 AGAGTGACATTTTGTCTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4024
Qy 781 TTAACCTTTAAATACCCCAATGTATCCCAATACATCATTTCAAAAGTGTAAATTAATAAAA 840
Db 4025 TTAACCTTTAAATACCCCAATGTATCCCAATACATCATTTCAAAAGTGTAAATTAATAAAA 4084
Qy 841 ACAATTAATGATGAGTACTTTTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 900
Db 4085 ACAATTAATGATGAGTACTTTTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4144
Qy 901 GAGTATATATTTATGCTGACAGACATCTCAATTTGGACTAGTACATTTTCAGGTGCTC 960
Db 4145 GAGTATATATTTATGCTGACAGACATCTCAATTTGGACTAGTACATTTTCAGGTGCTC 4204
Qy 961 AGTAGCCACATGTGCTGAGCAGTACTGTTATTTGATGGCAC 1001
Db 4205 AGTAGCCACATGTGCTGAGCAGTACTGTTATTTGATGGCAC 4245

RESULT 2

US-09-949-016-63399/c

; Sequence 63399, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 63399

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-63399

Query Match 60.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 3.2e-148;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 148 TAAGTCAGAAAGGTCGGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCG 207
Db 601 TAAGTCAGAAAGGTCGGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCG 542
Qy 208 CCGGTCACAAAGCAGCTAGAGACTCTCTCCAGCTGAGGACACCTCTCTTTATTTCAGCAGA 267
Db 541 CCGGTCACAAAGCAGCTAGAGACTCTCTCCAGCTGAGGACACCTCTCTTTATTTCAGCAGA 482
Qy 268 TACACACTGAGTCCCAACTCGGTAAATGAGCGTTCGCAAAATTCGAGAAATTCAGCAAT 327
Db 481 TACACACTGAGTCCCAACTCGGTAAATGAGCGTTCGCAAAATTCGAGAAATTCAGCAAT 422
Qy 328 TGCCAGACAGTCAGGACCCCTGTTCTCAGAGACTCATACCCCTAGAGTAGTGGTGTTTA 387
Db 421 TGCCAGACAGTCAGGACCCCTGTTCTCAGAGACTCATACCCCTAGAGTAGTGGTGTTTA 362
Qy 388 GTAGAAATTAATGCTGAGCTGCTTATGTCAATTTCCAGTTTTTTTAGTAGCCACATTTAAACA 447
Db 361 GTAGAAATTAATGCTGAGCTGCTTATGTCAATTTCCAGTTTTTTTAGTAGCCACATTTAAACA 302
Qy 448 GGTAAAAAGGCTGGCGCAGTCGCTCACACCTGTAAATCCAGACATTTGGAGGCTGAG 507
Db 301 RGTAATAAAGGCTGGCGCAGTCGCTCACACCTGTAAATCCAGACATTTGGAGGCTGAG 242
Qy 508 GCAGGACATCATTGTTGTCAGAGTTTGAGACTAGCTGGCCCAACATGGCGAACTCT 567
Db 241 GCAGGACATCATTGTTGTCAGAGTTTGAGACTAGCTGGCCCAACATGGCGAACTCT 182
Qy 568 GTCTCTAAAAAATAACAAAAATTAAGCCTGGCATGGTGGCGGCGCTGTAAATCTCAGC 627
Db 181 GTCTCTAAAAAATAACAAAAATTAAGCCTGGCATGGTGGCGGCGCTGTAAATCTCAGC 122
Qy 628 TGCTCAGGAGCGCCAGACACAAGAAATCACTTAAACCAGAGGTGGAGTTGCAGTGAGC 687
Db 121 TGCTCAGGAGCGCCAGACACAAGAAATCACTTAAACCAGAGGTGGAGTTGCAGTGAGC 62
Qy 688 TGAGATGTCGCACTCACTCAACCTGGGAGACAGATGACACTTTTGTCTCAAAAGAA 747
Db 61 TGAGATGTCGCACTCACTCAACCTGGGAGACAGATGACACTTTTGTCTCAAAAGAA 2
Qy 748 A 748
Db 1 A 1

RESULT 3

US-09-949-016-63400/c

; Sequence 63400, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 63400

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-63400

Query Match	60.0%;	Score 600.6;	DB 4;	Length 601;
Best Local Similarity	99.8%;	Pred. No. 3.2e-148;		
Matches 600;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY	167	GGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGCTGACAAAGCAGGTAG	226	
DB	601	GGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGCTGACAAAGCAGGTAG	542	
QY	227	AGACTCTCGCAGGTGAGGACACCTCTCTTTATTACGACAGATACACACTGAGTGCCAACT	286	
DB	541	AGACTCTCGCAGGTGAGGACACCTCTCTTTATTACGACAGATACACACTGAGTGCCAACT	482	
QY	287	CGGTAACTATGGAGCGTTTGCCAAATCTGAGAAATCCAGCAATGGCCAAACACACTCAGGACC	346	
DB	481	CGGTAACTATGGAGCGTTTGCCAAATCTGAGAAATCCAGCAATGGCCAAACACACTCAGGACC	422	
QY	347	CCTGTTCTCA CAGAGCTCATACCCCTAGAGTAGTGGTGTTTAGTAGAAAAATATGCTGAGCT	406	
DB	421	CCTGTTCTCA CAGAGCTCATACCCCTAGAGTAGTGGTGTTTAGTAGAAAAATATGCTGAGCT	362	
QY	407	GCTTATGTCATTTCCAGCTTTTTTATAGTAGCCACATTTAAACAGGTAAACAGCGCTGGCGC	466	
DB	361	GCTTATGTCATTTCCAGCTTTTTTATAGTAGCCACATTTAAACAGGTAAACAGCGCTGGCGC	302	
QY	467	AGTGGCTCACACCTGTAAATCCAGCACTTTGGGAGGCTGAGGCAGGCAGATCACCTTTGG	526	
DB	301	AGTGGCTCACACCTGTAAATCCAGCACTTTGGGAGGCTGAGGCAGGCAGATCACCTTTGG	242	
QY	527	TCAGGAGTTTGAGACTAGCCTGGGCCAACATGGCGAAACTCTGTCTCTAAAAAANAATACA	586	
DB	241	TCAGGAGTTTGAGACTAGCCTGGGCCAACATGGCGAAACTCTGTCTCTAAAAAANAATACA	182	
QY	587	AAATTAGCCTGGCATGGTGGCGGGCGCTCTAATCTCAGCTGCTCAGGAGCCGAGACA	646	
DB	181	AAATTAGCCTGGCATGGTGGCGGGCGCTCTAATCTCAGCTGCTCAGGAGCCGAGACA	122	
QY	647	CAGAAATCACTTAAACCCAGAGGTGGAGGTTGCAGTGAGCTGAGATCGTGCCACTCACT	706	
DB	121	CAGAAATCACTTAAACCCAGAGGTGGAGGTTGCAGTGAGCTGAGATCGTGCCACTCACT	62	
QY	707	CCAACCTGGGAGACAGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAA	766	
DB	61	CCAACCTGGGAGACAGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAA	2	
QY	767	A 767		
DB	1	A 1		

RESULT 4
US-09-949-016-63401/c
; Sequence 63401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63401
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-63401			
Query Match	60.0%; Score 600.6; DB 4; Length 601;		
Best Local Similarity	99.8%; Pred. No. 3.2e-148;		
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	242	GAGGACACCTCTCTTTATTGACGAGATACACACTGAGTGCCTCGTAACATGGAGCG	301
Db	601	GAGGACACCTCTCTTTATTGACGAGATACACACTGAGTGCCTCGTAACATGGAGCG	542
Qy	302	TTGCCAANTTCTGAGAAATCCAGCAATTTGCCAAGACAGTCAAGGACCCCTGTTCTCACAGAG	361
Db	541	TTGCCAANTTCTGAGAAATCCAGCAATTTGCCAAGACAGTCAAGGACCCCTGTTCTCACAGAG	482
Qy	362	CTCATACCCCTAGAGTAGTGGTGTTTAGTAGAAATAATGCTGAGTGTCTTATGTCATTTC	421
Db	481	CTCATACCCCTAGAGTAGTGGTGTTTAGTAGAAATAATGCTGAGTGTCTTATGTCATTTC	422
Qy	422	AGTTTTTTTAGTGCACATTTAAACAGGTAAAAAGGCTGGGCGCAGTGGCTCACACCTG	481
Db	421	AGTTTTTTTAGTGCACATTTAAACAGGTAAAAAGGCTGGGCGCAGTGGCTCACACCTG	362
Qy	482	TAATCCCGACACTTTGGGAGGCTGAGCGAGCAGATCACCTTTGGTCAGGAGTTTGAGAC	541
Db	361	TAATCCCGACACTTTGGGAGGCTGAGCGAGCAGATCACCTTTGGTCAGGAGTTTGAGAC	302
Qy	542	TAGCTGGCCAAACATGGCGAAAATCTCTCTCTAAAAAAAATAACAAAAATTTAGCCTCGCA	601
Db	301	YAGCTGGCCAAACATGGCGAAAATCTCTCTCTAAAAAAAATAACAAAAATTTAGCCTCGCA	242
Qy	602	TGTTGGCGGCGCCTGTAAATCTCAGCTGTCTCAGGAGCCGAGACACAGAATCACTTAAA	661
Db	241	TGTTGGCGGCGCCTGTAAATCTCAGCTGTCTCAGGAGCCGAGACACAGAATCACTTAAA	182
Qy	662	CCGAGGAGTGGAGGTTGAGTGAGTGCCTCCTCACTCTCAACCTGGGAGACA	721
Db	181	CCCAGGAGTGGAGGTTGAGTGAGTGCCTCCTCACTCTCAACCTGGGAGACA	122
Qy	722	GAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAATAAAGAACAGGTGAAGT	781
Db	121	GAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAATAAAGAACAGGTGAAGT	62
Qy	782	TAACCTTAAATACCCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATATATAA	841
Db	61	TAACCTTAAATACCCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATATATAA	2
Qy	842	C 842	
Db		1 C 1	

RESULT 5
US-09-949-016-63402/c
; Sequence 63402, Application US/09949016
; Patent No. 6812139
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63402
; LENGTH: 601
; TYPE: DNA

; ORGANISM: Human
US-09-949-016-63402

Query Match 54.5%; Score 546; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 7.1e-134; Indels 0; Gaps 0;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 456 AGGCTGGCGCAGTGGCTCACCTGTAATCCAGCACTTTGGAGGCTGAGGCGAG 515
Db 601 AGGCTGGCGCAGTGGCTCACCTGTAATCCAGCACTTTGGAGGCTGAGGCGAG 542
Qy 516 ATCACTTTGGTCAGGAGTTGAGACTAGCTGCGCAACATGGCGAACTGTCTCTAA 575
Db 541 ATCACTTTGGTCAGGAGTTGAGACTAGCTGCGCAACATGGCGAACTGTCTCTAA 482
Qy 576 AAAAAAATACAAAAATTAGCTGCGCATGGTGGCGCGCCCTGTAATCTCAGCTGCTCAGG 635
Db 481 AAAAAAATACAAAAATTAGCTGCGCATGGTGGCGCGCCCTGTAATCTCAGCTGCTCAGG 422
Qy 636 AGGCGGAGACACAAAGATCACTTAAACCCAGGAGGTGGAGTTGCAGTGGAGTGC 695
Db 421 AGGCGGAGACACAAAGATCACTTAAACCCAGGAGGTGGAGTTGCAGTGGAGTGC 362
Qy 696 TGCACACTCACTCCAACTGGGAGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAAA 755
Db 361 TGCACACTCACTCCAACTGGGAGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAAA 302
Qy 756 ACAAGTAAAAAGAAACAGAGTGAAGTTAACTTTTAATACCAATGATCCCAATACAAT 815
Db 301 ACAAGTAAAAAGAAACAGAGTGAAGTTAACTTTTAATACCAATGATCCCAATACAAT 242
Qy 816 CATTTCAAAGTGAATTAATATAAACAATATGAATGAGATCTTTTACATCTTTTCTT 875
Db 241 CATTTCAAAGTGAATTAATATAAACAATATGAATGAGATCTTTTACATCTTTTCTT 182
Qy 876 GTTTTCATATTAAGCTTTGAAAGTGAATATATATGATGATGCTGACGACATCTCAATT 935
Db 181 GTTTTCATATTAAGCTTTGAAAGTGAATATATATGATGATGCTGACGACATCTCAATT 122
Qy 936 TGGACTAGCTACATTTCCAGGTGCTCAGTAGCCACATGGCTAGCAGTACTGATTTGGA 995
Db 121 TGGACTAGCTACATTTCCAGGTGCTCAGTAGCCACATGGCTAGCAGTACTGATTTGGA 62
Qy 996 TGGCAC 1001
Db 61 TGGCAC 56
```

RESULT 6

US-09-949-016-63403/c
; Sequence 63403, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63403
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63403

Query Match 25.4%; Score 254; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.9e-57;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 748 AAAAAAACAAGTAAAGAAACAGGTGAAGTTAACTTTAAATACCCCAATGATCCCA 807
Db 601 AAAAAAACAAGTAAAGAAACAGGTGAAGTTAACTTTAAATACCCCAATGATCCCA 542
Qy 808 AATACAAATCAATTTCAAAGTGAATTAATATAAACAATTAATGAATGAGATCTTTACATT 867
Db 541 AATACAAATCAATTTCAAAGTGAATTAATATAAACAATTAATGAATGAGATCTTTACATT 482
Qy 868 CTTTTCTTTGTTTCATATTAAGTCTTTGAAAGTGAATATATGTTATGCTGACAGACA 927
Db 481 CTTTTCTTTGTTTCATATTAAGTCTTTGAAAGTGAATATATGTTATGCTGACAGACA 422
Qy 928 TCTCAATTTGGACTAGCTACATTTTCAGGTGCTCAGTAGCCACATGTCGCTAGCAGTTACT 987
Db 421 TCTCAATTTGGACTAGCTACATTTTCAGGTGCTCAGTAGCCACATGTCGCTAGCAGTTACT 362
Qy 988 GTATTGGATGGCAC 1001
Db 361 GTATTGGATGGCAC 348
```

RESULT 7

US-09-813-133A-3/c
; Sequence 3, Application US/09813133A
; Patent No. 6455294
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USFS THEREOF
; FILE REFERENCE: CL001173
; CURRENT APPLICATION NUMBER: US/09/813,133A
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Human
US-09-813-133A-3

Query Match 24.0%; Score 240.2; DB 3; Length 55827;
Best Local Similarity 68.6%; Pred. No. 1.2e-52;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

```
Qy 363 TCATACCCCTAGAGTAGTGGTGTGTTTAGTGAATAAATGCTGAGCTGCTTATGTCATTTCCA 422
Db 46943 TGAATCCTCAGACACAGGGTTGTTTAAAGAAATATATGCAAGTCACATATTTACTTTAA 46884
Qy 423 GTTTTATAGTACGACATTAACACAGTAAAA--AAGGCTGGGCGCAGTGGCTCACACCT 480
Db 46883 ATTTTCAAGTTACACCTTTTAAAGAAAAATAAAGGCGGAGGCGGCTGCTCACACCT 46824
Qy 481 GTATCCAGCACTTTGGGAGGCTGAGGCGGAGATCACCTTTGGTCAGAGTTTGAGA 540
Db 46823 ATAATCCAGCACTTTGGGAGGCTGAGGCGGAGATCACCTTTGAGATCAGAGTTTGAGA 46764
Qy 541 CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAATAACAAAAATAGCCTGGC 600
Db 46763 CCAGCGAGGCCAACATGGTGAACCCCGTCTCTACCAAAA--TACAAAAATAGCCAGGC 46706
Qy 601 ATGGTGGCGGCGCTGTATCTCAGCTGCTCAGGAGGCGGAGACACAAGAAATCACTTAA 660
Db 46705 ATAGTGGTCACGCTGTAATCCAGCTACTTTGGAGGCTGAGGCGGAGAACTTGCCTGA 46646
Qy 661 ACCCAGGAGTGGAGGTTGCAGTGAAGTCTGAGATCGTCCACT--CACTCCACCTGGGAGA 719
Db 46645 ACCCAGGAGGAGAGGTTGCAGTGAAGTCTGAGATCATGACACTGCACTCCAGCCTGGATGA 46586
```


Db 64131 AACCCGGAGGCGAGGTTGTCAGTGAACCGAGATCGGCCCACTGCACTCTAGCTGGTG 64072
Qy 719 ACAGAGTGACACTTT-TGTCTCAAAAGAGAAAAAACAAGTAAABAAACAGGTG 777
Db 64071 ACAGAGAGAGACTACATCTCAACAACAACAACAACAATAAATAATATAAGAGA 64012
Qy 778 AAGTTAACTTTAAATAACCCCAATGTATCCAAATACAAATCAATTTCAAAAGTGTAAATTAAT 837
Db 64011 TAGGTAGAAATGATTTTATGTTTAT--CTAACCAATATATCAAAATATTTCAATTCA 63954
Qy 838 AAAACAATATGAATGAGATCTTACATCTCTTTTCTTTTCTTTTCAATTAAGCTTTTGAA 897
Db 63953 ACATGTAATCAATATAAAAAAATAAGATACTTTATATCTTTTTCACACTAAGATTGGA 63894
Qy 898 AGTGAGTATATGTTATGTCGACGACACATCTCAATTTGGACTAGCTACATTTTCAGGTG 957
Db 63893 AATGGTATATATTTTACACTTATAGCATATCTCAATTCAGAATAGCCTCATAGTGGTG 63834
Qy 958 CTCA 961
Db 63833 GGCA 63830

RESULT 10
US-09-949-016-15779/c
; Sequence 15779, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15779
; LENGTH: 265038
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(265038)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15779

Query Match 23.0%; Score 230.2; DB 4; Length 265038;
Best Local Similarity 69.1%; Pred. No. 1e-49;
Matches 375; Conservative 0; Mismatches 158; Indels 10; Gaps 4;
Qy 437 ACATTAAACAGGTAAAAAAGCGTGGCGGAGTGGCTCACACCTGTAAATCCAGCACTTT 496
Db 52178 AAAACAAAAAATAATTCAGGCTAGGTGGCGTGGCTCACACCTATAATCTCAGCACTTT 52119
Qy 497 GGGAGGCTGAGCGAGGCGAGATCACCTTTGGTCAGGAGTTTGGAGACTAGCCTGGGCCAACAT 556
Db 52118 GGGAGGCGCAGGCGAGGAGATTAAGTGGTGGGAGTTTGGAGCAGCCTGGGCCAACAT 52059
Qy 557 GCGGAAACTCTGTCTCTTAAAAAATAAATAAATTTAGCCTGGCATGGTGGCGGCGCCT 616
Db 52058 GGTGAAACCCCATCTCT-ACTAAAAATACAAAAATTAACAGGTTGGTGGTGCATGCCT 52000
Qy 617 GTAATCTCAGTGTCTCAGGAGCGGAGACACAAGAAATCACTTAACCCAGAGGTGGAGG 676
Db 51999 GTAATCCAGCTGTCTCAGGAGGCTGAGGCAACAAGAAATTTGCTTGAATCCAGGAGCAGAGT 51940

Qy 677 TTCAGTGTAGCTGAGATCGTGCCA-CTCACTCCAACTCGGAGACAGAGTGACACTTTTG 735
Db 51939 TTCTAGTGTAGCCCAAGATCATGCCCACACACTCCAGCTCGGCCAACAGACTGAGACTCTGT 51880
Qy 736 TCTCAAAAAGAAAAAACAACAA--GTAAAAAGAAACAGGTGAAGTTAACTTTAATAA 793
Db 51879 CTCAAAAATAATAATAACTAATAAGGGGAAAAAAGAAACAGGTGAATAATCATTTGTAATAT 51820
Qy 794 CCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATTAATAAACAATTAATGAATG 853
Db 51819 GTCTTATTAAACCGAATCTGCTACACATTAATTTCAAACGTAATCGATCTAGGTTTATT 51760
Qy 854 AGATACTTTACATCTCTTTCTGTTTTCATATTAAGTCTTTTGAAGAGTATATATGTT 913
Db 51759 AAG-----ACAGTTTCCCCTTGTCTTCTACTAAGCCTGGGAAAACTGCTATCGGTTCC 51706
Qy 914 ATCTGACAGCACATCTCAATTTGGACTAGCTACATTTCAAGTGTCTCAGTAGCCACATGT 973
Db 51705 AACTCACAGTATGGCTCAGGGCGAACCCAGCCACATTTTCGGGTGCTTAGGGGCCACATAC 51646
Qy 974 GGC 976
Db 51645 TGC 51643

RESULT 11
US-09-949-016-12893/c
; Sequence 12893, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12893
; LENGTH: 131724
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12893

Query Match 23.0%; Score 229.8; DB 4; Length 131724;
Best Local Similarity 79.6%; Pred. No. 9.5e-50;
Matches 284; Conservative 0; Mismatches 72; Indels 1; Gaps 1;
Qy 436 CACATTAAAAACAGGTAAAAAGCGTGGCGGAGTGGCTCACACCTGTAAATCCAGCACTT 495
Db 5353 CAATTAATAAACACAACTACAGCGCGGCGAGTGGCTCACGCTGTAAATCCAGCACTC 5294
Qy 496 TGGAGGCTGAGGCGAGGAGATCACCTTTGGTCAGGAGTTTGGAGACTAGCCTGGCCACA 555
Db 5293 TGGGAAGCCGAGCGGCTGGATCACCTGAGTTCAGGAGTTTGGAGACCAGCCTGGCCAATA 5234
Qy 556 TGGCGAAACTCTGTCTCTAAAAAATAAATAAATTTAGCCTGCGATGGTGGCGGCGCC 615
Db 5233 TGGTGAATCCCGTCTCTAACTAAATAAATAAATAAATTAACCGGTGTGGTGGCGTGTGCC 5174
Qy 616 TGTAAATCTAGCTGCTCAGGAGCGGAGACACAAGAAATCACTTAAACCCAGAGGTGGAG 675
Db 5173 TGTAGTCCCAGCTACTCGGAGGCGGAGGAGGAGAGAAATCGCTTGAAACCCAGAGGCGGAG 5114
Qy 676 GTTGCACTGAGCTGAGATCGTGCACCT-CACTCCAACTGGGAGACAGAGTGACACTTTT 734
Db 5113 GTTGCACTGAGCTGAGATCATGCCACTACAGCCTTGGGGGACAGAGGAGACTCCT 5054

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 400.727 Seconds

(without alignments)
15667.668 Million cell updates/sec

Title: US-09-936-271c-13_COPY_6000_7000

Perfect score: 1001

Sequence: 1 gacaacaaagagccccaag.....gttactgtatggatggcac 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252.6	25.2	52216	9	US-09-747-810-1
2	248.8	24.9	56098	13	US-10-087-192-1612
3	240.2	24.0	55827	9	US-09-813-133A-3
4	240.2	24.0	55827	14	US-10-212-877-3
5	240.2	24.0	55827	21	US-10-954-110-3
6	235.6	23.5	27189	13	US-10-087-192-1828
7	235.6	23.5	84105	19	US-10-741-601-5637

c	8	235.2	23.5	1980090	20	US-10-719-993-6815	Sequence 6815, Ap
c	9	235.2	23.5	1980090	21	US-10-741-600-17676	Sequence 17676, A
c	10	227.4	22.7	23071	9	US-09-764-864-1673	Sequence 1673, Ap
c	11	226.2	22.6	133955	13	US-10-087-192-1984	Sequence 1984, Ap
c	12	225.4	22.5	74424	21	US-10-773-678-153	Sequence 153, App
c	13	225.4	22.5	325791	11	US-09-768-185A-1	Sequence 1, Appli
c	14	225.2	22.5	35641	9	US-09-962-436-306	Sequence 306, App
c	15	225.2	22.5	35641	9	US-09-880-107-2225	Sequence 2225, Ap
c	16	225.2	22.5	35641	21	US-10-843-841A-2765	Sequence 2765, Ap
c	17	224.8	22.5	629	13	US-10-027-632-252422	Sequence 252422,
c	18	224.8	22.5	629	13	US-10-027-632-252423	Sequence 252423,
c	19	224.8	22.5	629	17	US-10-027-632-252422	Sequence 252422,
c	20	224.8	22.5	629	17	US-10-027-632-252423	Sequence 252423,
c	21	224.8	22.5	629	19	US-10-681-818-155	Sequence 155, App
c	22	224.8	22.5	629	19	US-10-681-818-156	Sequence 156, App
c	23	224.6	22.4	653122	13	US-10-087-192-226	Sequence 226, App
c	24	223.2	22.3	67253	22	US-10-737-082-88	Sequence 88, Appl
c	25	223.2	22.3	67253	22	US-10-765-790-88	Sequence 88, Appl
c	26	221.6	22.1	2591	19	US-10-755-889-83	Sequence 83, Appl
c	27	221.2	22.1	84409	19	US-10-741-601-5696	Sequence 5696, Ap
c	28	221.2	22.1	84409	21	US-10-741-600-17771	Sequence 17771, A
c	29	220.8	22.1	96587	11	US-09-997-722-250	Sequence 250, App
c	30	220.6	22.0	203264	13	US-10-087-192-988	Sequence 988, App
c	31	220.2	22.0	126990	19	US-10-717-597-68	Sequence 58, Appl
c	32	220	22.0	32146	9	US-09-764-860-797	Sequence 797, App
c	33	220	22.0	32146	14	US-10-074-095-797	Sequence 797, App
c	34	220	22.0	32146	17	US-10-212-872-797	Sequence 797, App
c	35	220	22.0	32248	9	US-09-764-860-802	Sequence 802, App
c	36	220	22.0	32248	14	US-10-074-095-802	Sequence 802, App
c	37	220	22.0	32248	17	US-10-212-872-802	Sequence 802, App
c	38	220	22.0	96589	18	US-10-052-482-214	Sequence 214, App
c	39	219.2	21.9	100267	18	US-10-240-425-1470	Sequence 1470, Ap
c	40	218.6	21.8	4329	10	US-09-764-891-8960	Sequence 8960, Ap
c	41	218.6	21.8	43159	21	US-10-741-600-17953	Sequence 17953, A
c	42	218.6	21.8	106236	21	US-10-741-600-17759	Sequence 17759, A
c	43	218.6	21.8	366803	20	US-10-719-993-6805	Sequence 6805, Ap
c	44	218.6	21.8	410846	21	US-10-481-613-1	Sequence 1, Appli
c	45	218.4	21.8	2312	9	US-09-764-877-3606	Sequence 3606, Ap

ALIGNMENTS

RESULT 1
US-09-747-810-1
; Sequence 1, Application US/09747810
; Patent No. US20020012903A1
; GENERAL INFORMATION:
; APPLICANT: Goldman, Steven A.
; APPLICANT: Okano, Hideyuki
; TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGENITOR CELLS
; FILE REFERENCE: 19603/3580
; CURRENT APPLICATION NUMBER: US/09/747,810
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,003
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-810-1

Query Match	25.2%	Score 252.6;	DB 9;	Length 52216;
Best Local Similarity	65.1%	Pred. No. 9.8e-49;		
Matches 403;	Conservative 0;	Mismatches 174;	Indels 6;	Gaps 4;
Qy	417	TTTCCAGTTTCTTAGTACCCACATTAAACAGGTAAACAGGTGGGCGGCTGCTCAC	476	
Db	41515	TTAAATTTCTTAGTACCCACATTAA--NAGACATTAAGCCGGCGGCTGCTCAC	41572	
Qy	477	ACCTGTAATCCACACCTTTGGGAGGCTGAGGAGGAGATCACCTTTGTCAGGAGTTT	536	

Db 538603 ATAGTCTATGTTCTTTTTCCTAAG-----TCTGAAATTCGGTGTGTAT-TTAC 538756
Qy 916 GCTGACAGCACATCTCAATTTGGACTAGTACATTTTCAGGTGCTCAGTAGCCACATGGG 975
Db 538755 ACTGACAGCACTCAATTTAGACTAGCCACATTTTCAGTGTGCTGATGTATC 538696
Qy 976 CTAGCAGTACTGTTATTTGGATGGCAC 1001
Db 538695 CTGCTGGCTACCCACACTGGACAGCAC 538670

RESULT 10
US-09-764-864-1673/c
; Sequence 1673, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1673
; LENGTH: 23071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1673

Query Match 22.7%; Score 227.4; DB 9; Length 23071;
Best Local Similarity 70.1%; Pred. No. 6.3e-43;
Matches 445; Conservative 0; Mismatches 156; Indels 34; Gaps 9;
Qy 381 GTGTTTAGTGAATAATAGTGTAGCTGTTATGTCATTTTCAGTTTTTTAGTACCCACAT 440
Db 10748 GTGCTAAGAGAAATATAATGTGAGCCACATATATAATTTTAAATTTTCCTGGCTCCACAT 10689
Qy 441 TAAACAG-GTAAAAAGGCTGGCGAGTGGCTCACACCTGTAAATCCCGACATTTGGG 499
Db 10688 AACAAAAGTGAAGAGGCTGGCGGCTGCTGCTGTAAATCCCGACATTTGGG 10629
Qy 500 AGGCTGAGGCGAGGATCACCTTTTGGTCAGGAGTTTGAGACTAGCCTGGGCCAACATGGC 559
Db 10628 AGGCTGAGAGCGGATCACCTGAGTGGGAGTTCGAGACTAGCTTAACCAACATGGA 10569
Qy 560 GAAACTCTGTCT---CTAAAAAAATACAAAAATTTAGCCTGGCATGGTGGCGGCGCCT 616
Db 10568 GAAATCCCATCTCTACTAAAAAAATTTACAAAAATTTAGCTGGGCGTGGTGACATGCCT 10509
Qy 617 GTATCTCAGCTGCTCAGGAGCGGAGACACAAGAACTTAAACCCAGAGGTGGAGG 676
Db 10508 GTATCCAGCTACTAGGAGGCTGAGGAGGAAATCGCTTTGAACCCGGAGGCGGAGG 10449
Qy 677 TTGCAGTGAGCTGAGATCGTGCCACT-CACCTCCAACTGGGAGAGAGAGTGACACTTTTG 735
Db 10448 TTGCGGTGAGCGGAGATTGCGCCATTGCACTCCAGCTGGGCAACAAGCGAACTCCA 10389
Qy 736 TCTCAAAAGAAAAAACAAGTAA-----AAAAAACAAGTAAAGTTAACTTT 788
Db 10388 TCTCAAAAAAAACAAAAAACAACAAAAAAGTGAAGAAAGGAAACAGGGAAGTTAAATTA 10329
Qy 789 AA-----TAACCCAAATGATCCCAATACATCAATTTCAAGTGTAAATATAAACA 843
Db 10328 AATAGTTTAAACCAAGTGTATCCATGATATTAATTTCAACATGTAATCAACATTA--A 10271
Qy 844 ATTATGATGAGATACTTTTACATTTCTTTCTTTGTTTTCATATTAAAGTCTTTTGAAGTGAG 903
Db 10270 ACTATTACTGAGACATTTTACATTTGTT-----TATACATCTTCAAAATCCAT 10223
Qy 904 TATATATGTTATGCTG-ACAGCAATCTCAATTTGGAGTAGCTACATTTAGGTGCTCAG 962
Db 10222 TGTGTATTTTCTACTTTCACGACAACTCAATTTGGG--CAGCCACATTTCAAAATTTCTCGA 10165

Qy 963 TAGCCACATGTGGCTAGCAGTTACTGTATTTGGATG 997
Db 10164 TAGCCACAAATGCTATTGGCTATTTATTTGTTG 10130

RESULT 11
US-10-087-192-1984
; Sequence 1984, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1984
; LENGTH: 133955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1984

Query Match 22.6%; Score 226.2; DB 13; Length 133955;
Best Local Similarity 68.4%; Pred. No. 2.6e-42;
Matches 357; Conservative 0; Mismatches 158; Indels 7; Gaps 3;
Qy 425 TTTTGTAGTAGCCACATTAACAGGTAA-----AAAAGGCTGGCGCAGTGGCTCACACC 479
Db 115167 TTTCTAGTGGCCACATTTTAAAAACATAAGAAAGAACAGGCGCGTGGTGCCTCATGCC 115226
Qy 480 TGTAAATCCAGCAGCTTTGGGAGCTGAGCGAGGAGATCACCTTTGCTCAGGAGTTTGAG 539
Db 115227 TGTAAATCCAGCAGCTTTGGAGGCTGAGGCAAGCAATCACTTTGAGGCCAGGAGTTCAAG 115286
Qy 540 ACTAGCCTGGCCAACTATGGCGAAACTCTGTCTCTTAAAAAATAAATAAATAAATAAAGTGG 599
Db 115287 ACCAGCCTGTCAACATGGTGAAACTCTGTCTCT-CTAAAAATAACAAAAATTAGCCAGG 115345
Qy 600 CATGTGGCGGGCGCCTGTAACTCTCAGCTGCTCAGGAGGCGGAGACACAGAAATCACTTA 659
Db 115346 CATGTGGCGCACACCTGTAAATCCAGCTACTCAAGTAGGTGAGGCAACAAGAAATCACTTG 115405
Qy 660 AACCCAGAGGTGGAGGTTGCGAGTGCAGTGCAGATCGTGCCACT-CACTCCAACTGGGAG 718
Db 115406 AACCCGGAAGCGAGGTTGCGAGTGCAGGATCAGCCACTGTACTCTCAGCCTGTGTG 115465
Qy 719 ACAGAGTGACACTTTTGTCTCTCAAAAGAAAAAATAAAGTAAAAAAGAAAAACAGGTGA 778
Db 115466 ACAGAGCAAGACTCTGTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 115525
Qy 779 AGTTAACTTTAAATAACCAATGTATCCCAATACATCAATTTCAAGTGTAAATATA 838
Db 115526 ATTATTTTAACTAACTTTTATTTAAACCCAGTATCAATTTTCAAAATG 115585
Qy 839 AAACAATTTAATGATGAGATACTTTTACATTTCTTTTGTGTTTCTTATTTAACTTTTGA 898
Db 115586 CAATTCAATATAATTAATTTAGATATTTTACATTTCTTTTATGCTATGCTTTTGACA 115645
Qy 899 GTGAGTATATATGTTATGCTGACAGCACATCTCAATTTGGAC 940
Db 115646 TCATGTGTGTGTTTACATTTACATTTACAGCCCACTTTTAAATTCGGAC 115687

RESULT 12
US-10-773-678-153/c

; Sequence 153, Application US/10773678
; Publication No. US20050074879A1
; GENERAL INFORMATION:
; APPLICANT: Kairas, James G
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3
; FILE REFERENCE: ISPH-0828
; CURRENT APPLICATION NUMBER: US/10/773,678
; PRIOR FILING DATE: 2004-02-06
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 09/758,881
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: PCT/US00/09054
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/288,461
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 402
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 153
; LENGTH: 74424
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; OTHER INFORMATION: Antisense oligonucleotide
US-10-773-678-153

Query Match 22.5%; Score 225.4; DB 21; Length 74424;
Best Local Similarity 68.3%; Pred. No. 3.1e-42;
Matches 372; Conservative 0; Mismatches 166; Indels 7; Gaps 4;

Qy 340 CAGGACCCCTCTTCCACAGAGCTCATACCTAGAGTAGTGGTGTAGTGAATAATG 399
Db 38162 CAGGAGCTTTTACATTGAATAATAATACTAAGGAGCAGCATCCCAACAGAAAGAA 38103

Qy 400 CTGAGCTGCTTATGTCTATCTCCAGTTTTTTAGTAGCCACATTAAACAGGTAAAGAGGC 459
Db 38102 TACAGGCCACATGCTATTTAAATTTCTGTAGCCATCTTAAAGATTAAAGAGGC 38043

Qy 460 TGGCGCAGTGGCTCACACCTGTATATCCAGCATTTTGGGAGGCTGAGGCGAGGATCA 519
Db 38042 TGGGTGCGAGTGGCTCACGCTGTATCCAGAACTTTGGGAGGCGCAAGGCGAGGATCA 37983

Qy 520 CTTTGTGTCAGGAGTTTGAGACTAGCTGGCCACATGGCCAACTGTCTCTAAAGAA 579
Db 37982 --TGAGGTGAGGAGATCGAGACCATCTCTGGCCAAATATGGTGAAACCCCGTCTCT-ACTAA 37926

Qy 580 AAATACAAAAATTAGCCTGGCATGTGGCGGGCGCTGTATCTCAGCTGCTCAGGAGGC 639
Db 37925 AAATACAAAAATTAGCTGGGTGGTGCCACATGCTTTTATCCAGCTACTAGGAGGC 37866

Qy 640 CGAGACACAAGATCACTTAAACCCAGGAGGTGGAGGTTCAGTGAGCTGAGATCGTGCC 699
Db 37865 TGAGGCAGAGAAATCGCTTGAACCCAGGAGGAGGTTGCGAGTGAGCGAGATCGTGCC 37806

Qy 700 ACTCACTCAACCTGGGAGACAGAGTGCACCTTTTGTCTCAAAAGAGAAAGAAAGAA 759
Db 37805 ACTCACTCCAGCTGTATGACATAGCAGAGAC-TCCATCTCAAAAGAGAAAGAAAGAA 37747

Qy 760 GTAAAGAGAACAGGTGAAGTTAACTTTTATAACCCAACTATCCCAATACATCAATCATT 819
Db 37746 AAAGAAAGAACAGGTGAAGTTAACTTTTAAAT---GTATTTTATTTCAATTAATATATCC 37690

Qy 820 TCAAAGTGAATTAATAAACAATATGATGAGATCACTTTACATCTCTTTCTTTT 879
Db 37689 AAATGTTATTGCAACATATGACATACAAACATATTAATTAATTAATTAATTTT 37630

Qy 880 TCATA 884
Db 37629 GGAGA 37625

RESULT 13

Query Match

22.5%; Score 225.2; DB 9; Length 35641;

US-09-768-185A-1
; Sequence 1, Application US/09768185A
; Publication No. US20040185439A1
; GENERAL INFORMATION:
; APPLICANT: Cassel, Michael et al
; TITLE OF INVENTION: Estrogen receptor beta variants and
; FILE REFERENCE: CL000280
; CURRENT APPLICATION NUMBER: US/09/768,185A
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 09768185
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 325791
; TYPE: DNA
; ORGANISM: HUMAN
US-09-768-185A-1

Query Match 22.5%; Score 225.4; DB 11; Length 325791;
Best Local Similarity 82.6%; Pred. No. 5.8e-42;
Matches 270; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Qy 441 TAAACAGGTAAAGAGCTGGCGCAGTGGCTCACACCTCTAATCCAGCACCTTTGGGA 500
Db 282763 TGAAGAGAAACAGAGAGCGCGCGCAGTGGCTCACACCTCTAATCCAGCACCTTTGGGA 282822

Qy 501 GGCTGAGCGAGGAGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCTGGCCCAACATGGCG 560
Db 282823 GGCCAGCGAGGTGATCATCTGAGGTTCAGAGTTCAAGACTAGCTGGCCCAACATGGCA 282882

Qy 561 AAATCTGTCTTAAAGAGAAATACAAATTAAGCTGGCATGTGGCGGCGCTGTAA 620
Db 282883 AAATCTGTCTTAAAGAGAAATACAAATTAAGCTGGCATGTGGCGGCGCTGTAA 282942

Qy 621 TCTCAGCTGCTCAGGAGCGGAGACACAGAAATCACTTAACCCAGGAGTGGAGGTTC 680
Db 282943 TCCAGCTGCTCGGAGGCTGAAGCAGGAGATCGCTTGAACCCAGGAGGAGGTTC 283002

Qy 681 AGTGAGCTGAGATCGTGCCACT-CACCTCAACCTGGGAGACAGAGTGCACACTTTGTCTC 739
Db 283003 AGTGAGCTGAGATCGCACCACTGCACCTCCAGCTGGGCGCAGATCGAGACTCTGTCTCA 283062

Qy 740 AAAAG 766
Db 283063 AAAAG 283089

RESULT 14

US-09-962-436-306
; Sequence 306, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 306
; LENGTH: 35641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-306

Best Local Similarity 84.4%; Pred. No. 2.5e-42; Matches 265; Conservative 0; Mismatches 48; Indels 1; Gaps 1;	
Qy	455 AAGCTGGGGCAGTGGCTCACACCTGTAATCCAGCACTTTGGGAGGCTGAGGCAGCA 514
Db	14689 AAGCTGGGCTAGTGGCTCACGCTGTGTAATCCAGCACTTTGGGAGGCTGAGGCAGCC 14748
Qy	515 GATCACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCACATGGCGAACTCTGTCTCTA 574
Db	14749 GATCACCTGAGTTCAGGAGTTTGAGACCAAGCTTGGCCAAATATGGTGAACCCCGTCTCTA 14808
Qy	575 AAAAAAATACAAAAATTAGCCTGGCATGGTGGCGGGCGCCTGTAATCTCAGCTGCTCAG 634
Db	14809 TTAATAATACCAAAATTAGCCAGGGCGTGGTGGCGGCACCTGTAAATCCAGCTACTCAG 14868
Qy	635 GAGGCGCGAGACACAAGAATCACTTAAACCCAGGAGGTGGAGGTTCGAGTGAGCTGAGATC 694
Db	14869 GAGGCTGAGCGCAGGAGNAATCACTTTGAACCCAGGAGCGGAGGTTGCAGTGAGCTGAGATC 14928
Qy	695 GTGCCACT-CACTCCAACTGGGAGACAGAGTGACAATTTTGTCTCAAAAGAAAAAAA 753
Db	14929 ACACTACTGCACCTCCAGCCTTGGSCAACAGAGTGAGACTCGGTCTCAAAAAAAA 14988
Qy	754 AAACAAGTAAAAA 767
Db	14989 AAAAAACAAAAA 15002

Db	14869	GAGGCTGAGCGAGGAGATCCTTGAACCCAGGAGCGGAGGTTCCACTGTGAGCTGAGATC	14928
Qy	695	GTGCCACT..CACTCCAACTCTGGGAGACAGAGTGAACATTTTGTCTCAAAAAAGAAAAAAA	753
Db	14929	ACACTACTGACTCCAGCTCTGGCCACAGAGTGAGACTCGGTCTCAAAAAAGAAAAAAA	14988
Qy	754	AAACAAGTAAAAA	767
Db	14989	AAAAAACACAAAAA	15002

	Query Match	22.5%;	Score 225.2;	DB 9;	Length 35641;
	Best Local Similarity	84.4%;	Pred. No. 2.5e-42;		
	Matches 265;	Conservative 0;	Mismatches 48;	Indels 1;	Gaps 1;
Qy	455	AAGGCTGGGGCAGTGGCTCACACCTGTATATCCAGCACTTTGGGAGGTGAGGCAGGCA	514		
Db	14689	AAGGCTGGGCTCAGTGGCTCAGCTGTATATCCAGCACTTTGGGAGGCTGAGGCAGGCC	14748		
Qy	515	GATCACCTTGTGTGAGGAGTTTGAGACTAGCTGGCCACATGGCGAACTCTGTCTCTA	574		
Db	14749	GATCACCTTGAGGTGAGGAGTTTGAGACCACTGGCCATATGTGTAAACCCCGTCTCTA	14808		
Qy	575	AAAAAAATACAAAAATTAGCCTGGCATGTGTGGCGGCGCCTGTGTATCTCAAGCTGCTCAG	634		
Db	14809	TTAAAAATACAAAAAATTAGCCAGGGCGTGGTGTGGCGGCACTGTATCTCCAGCTACTCAG	14868		
Qy	635	GAGGCGCCAGACACAAGAATCACATTAACCCAGAGGTGGAGGTTTGCAGTGAGCTCAGATC	694		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2005, 00:19:06 ; Search time 2176.56 Seconds

(without alignments)
17505.766 Million cell updates/sec

Title: US-09-936-271C-13_COPY_6000_7000

Perfect score: 1001

Sequence: 1 gacaaccaagagcccccag.....gttactgtatggatggcac 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	226.8	22.7	908	6	CD245375 AGENCOURT
2	223	22.3	694	8	BZ603010 WHAAX35TF
3	221.6	22.1	1750	3	AF119908 Homo sapi
4	220	22.0	769	6	CD370173 UI-H-Ftl1
5	219.6	21.9	500	5	BX491983 DKFZp781O
6	218.4	21.8	837	2	BE790769 601581950
7	216	21.6	649	8	B71494 RPI111-8N9
8	216	21.6	1019	6	CD518470 AGENCOURT
9	215.6	21.5	2449	3	CR749224 Homo sapi
10	215.4	21.5	573	5	BX477553 DKFZp686H
11	215.4	21.5	823	6	CD101954 AGENCOURT
12	215	21.5	397	1	AV759632 AV759632
13	215	21.5	587	5	BX644875 DKFZp781D
14	215	21.5	633	7	CK905714 1j76h09.x
15	214	21.4	491	2	AW516097 xt60f05.x
16	214	21.4	4593	3	HSM803561
17	213.4	21.3	517	8	AQ264959 CITBI-El-
18	213	21.3	569	8	AQ490878 RPI11-2
19	212.8	21.3	700	5	BX508919 DKFZp686G
20	212.6	21.2	790	6	CD643827 AGENCOURT
21	212.4	21.2	619	5	BX485408 DKFZp686B
22	211.8	21.2	576	6	CD652305 AGENCOURT
23	211.8	21.2	731	7	CN307840 170004182
24	211.8	21.2	15970	8	AQ839852 260013-C5

C 25	211.4	21.1	537	8	AQ390925	CITBI-El-
26	211.4	21.1	624	9	AG019812	AG019812 Homo sapi
27	211.4	21.1	654	7	CK003802	AGENCOURT
28	211.4	21.1	687	6	CD237038	FNPAFH09
29	211.4	21.1	692	5	BU662309	CL84C05.2
30	211.2	21.1	338	2	AW502796	UI-HF-BR0
31	211.2	21.1	596	8	AQ320557	RPI111-96
32	211.2	21.1	801	8	BZ600703	WHASC76TF
33	211	21.1	552	8	B68316	CIT978SK-A-
34	210.8	21.1	657	7	CF123536	UI-HF-CHO
35	210.6	21.0	1016	5	BQ674146	AGENCOURT
36	210.4	21.0	388	2	AW069227	cr41h09.x
37	210.4	21.0	735	7	CF146929	UI-HF-CB0
38	210.2	21.0	671	9	AG079103	Pan trogl
39	210.2	21.0	849	6	CB991080	AGENCOURT
40	209.8	21.0	397	1	A1421950	tf45d05.x
41	209.8	21.0	463	2	AW026305	wv11d10.x
42	209.6	20.9	461	8	B65075	CIT-HSP-201
43	209.6	20.9	501	5	BX485916	DKFZp686E
44	209.6	20.9	604	5	BM990713	UI-H-DIO-
45	209.2	20.9	534	6	CB050438	NISC_gj17

ALIGNMENTS

RESULT 1
CD245375/c
LOCUS CD245375 908 bp mRNA linear EST 22-MAY-2003
DEFINITION AGENCOURT_14128276 NIH_MGC_181 Homo sapiens cdna clone
IMAGE:30374350 5', mRNA sequence.
ACCESSION CD245375
VERSION EST.
KEYWORDS CD245375.1 GI:31005839
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 908)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: NDAM433 row: j column: 23
High quality sequence start: 10
High quality sequence stop: 557.
Location/Qualifiers
1. 908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:30374350"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/note="Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 Kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 22.7%; Score 226.8; DB 6; Length 908;
Best Local Similarity 69.3%; Pred. No. 4.6e-25;

	Matches	384;	Conservative	0;	Mismatches	157;	Indels	13;	Gaps	5;
Qy	437	ACATTAAACAGGTGTA	AAAAAGGCTGGCGCAGTGGCTCACACCTGTGTAATCCCGACACTTT	496						
Db	680	AAATGTAAAGAAAAAAGT	CGCGCGGCGGCTCACGCTGTAAATCCCGACACTTT	621						
Qy	497	GGGAGGCTGAGGCGAGAT	CACCTTTGGTCAGGAGTTTGGAGACTAGCTGGCGCCACAT	556						
Db	620	GGGAGACCGAGGCGCGGAT	CAAC - GACGTCCGAGATCGAGACCATCTCTGTTAAACAC	563						
Qy	557	GGCGAAACTCTGTCTT	AAAAAAAATACAAAAAATTAGCTGGCATGGTGGCGGCGCCT	616						
Db	562	GGTGAACCTCGTCTT	ACTTAAATAACAAAAAATTAGCCAGGCATGGTGGCGGATGCCT	503						
Qy	617	GTAATCTCAGCTGCTCAGG	AGGCGGAGACACAGAATCACTTAAACCCAGGAGTGGAGG	676						
Db	502	GTAGTCCAGCTACTTGG	AGGCTGAGGCGAGGAGATGGCGTGAACCCGGGAGGCAGAGC	443						
Qy	677	TTCCAGTGAGCTGAGATCG	TGGCCACTCACCTCCACCTGGGAGACAGAGTGACACTTTTGT	736						
Db	442	TTGCAGTGAGCTGAGAT	GAGACCACTGCATCTCCACGAGCTGGGCGACAGAGAC - AGACT	384						
Qy	737	CTCAAAAGAAAAA	AAAAAACAAGTAAAAAGAAACAGGTGAAGTTAACTTTTAATAACCC	796						
Db	383	CTGTCTCAGAAAAA	AAAAAAGAAATAAGTCATAGATTTTAAATATATTAGCTC	324						
Qy	797	AATGTATCCCAATACAA	TCAATCTTCAAGTGTAATTAATAATAAACAAATATATGAATGAGA	856						
Db	323	AGTATGCTAAAAATAT	TATCATTTAA - - - - - CATGTAATCAACAAAGTTATTAAACAAA	270						
Qy	857	TACTTTACATTTCTTT	CTGTTTTCATATTAAGTCTTTGAAAGTGAGTATATATGT - TAT	915						
Db	269	TATTCTACATTCCTTT	TGG - - - - - GGGTAATAAATCTTTGAAATTAATAATGATTTTAT	213						
Qy	916	GCTGCACAGCACATCT	CAATTTGGACTAGCTACATTTCCAGGTGCTCAGTAGCCACATGTGG	975						
Db	212	ACTTACAGCTGATCT	TGATTTCAGACTAGCCACATTTCCAGTGCTCAGTAGCCCATGTGG	153						
Qy	976	CTAGCAGTTACTGT	989							
Db	152	CTAGCAACTATGTT	139							

RESULT 2
BZ603010
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BZ603010
WHAAX35TF Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-6P22, genomic survey sequence.
BZ603010
BZ603010.1 GI:31511472
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 694)
Volik, S., Zhao, S., Chin, K., Brebner, J. H., Herndon, D. R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J. W. and Collins, C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
<http://www.genomex.com>
Class: BAC ends.

```

FEATURES                                  Location/Qualifiers
  source                                  1..694
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="MCF7_1-6F22"
      /sex="female"
      /clone_lib="Human MCF7 breast cancer cell line library
      (MCF7_1)"
      /notes="Vector: pECBAC1; Site 1: HindIII; This library was
      constructed from MCF7 breast cancer cell line by Amplicon
      Express (http://www.genomex.com) using their standard
      procedure."

ORIGIN
Query Match                22.3%;   Score 223;   DB 8;   Length 694;
Best Local Similarity      79.1%;   Pred.No. 1.9e-24;
Matches 265; Conservative  0; Mismatches 70; Indels 0; Gaps 0;

Qy  439  ATTAACAACAGTAAAAAGCGCTGGGGCGCAGTGGCTCACACCTGTAATCCCAGACATTTGG 498
      |||||
Db    28  ACTAAAAATGGAATGTTCCTGGCGCGGTGCTCAGGCCTGTAATCCCAGACATTTGG 87

Qy  499  GAGCGTGAGCGCAGCATCACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCAAACATGG 558
      |||||
Db    88  GAGCGTGAGTTGGCGCAGATTGCCTGAGGTGAGAGTTCCAGACCAGCCTGGGCAACATGG 147

Qy  559  CGAAACTCTCTCTCTATAAAAAAATACAAAAATTAGCTGGCATGTTGGCGGCGCCTGT 618
      |||||
Db   148  CGAAACCTCTCTCTACTATAAATACAAAAATTAGCTGGCATGTTGGCGGCGCCTGT 207

Qy  619  AATCTCAGCTGCTCAGGAGGCCGAGACACAAAGAAATCACTTTAAACCCAGGAGGTGGAGGTT 678
      |||||
Db   208  AATCACAGCTACTCGGGAGGCTGAGGCAAGAGAATCGCTTGAGCCCGAGGAGCAGAGTT 267

Qy  679  GCAGTGAGCTGAGATCGTGGCACTCACTCCAACTGGGAGACAGAGTGACACTTTTGTCT 738
      |||||
Db   268  GCAGTGAGCGAGATGCGCCACTCGCCCTCCAGACTGGGTGACAGAGTAAGACTCTGTCTC 327

Qy  739  CAAAGAGAAAAAAGAAAAACAGTAAAAAGAAACA 773
      |||||
Db   328  AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 362

RESULT 3
AF119908                                1750 bp  mRNA  linear  HTC 08-MAY-2001
LOCUS
DEFINITION Homo sapiens PRO2955 mRNA, complete cds.
ACCESSION AF119908
VERSION AF119908.1 GI:7770252
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1750)
    Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
    and He,F.
  TITLE Functional prediction of the coding sequences of 79 new genes
    deduced by analysis of cDNA clones from human fetal liver
  JOURNAL Unpublished
  REFERENCE
    2 (bases 1 to 1750)
    Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
    and He,F.
  TITLE Direct Submission
  JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology,
    Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing,
    Beijing 100850, P. R. China
  JOURNAL
  REFERENCE
    Location/Qualifiers
FEATURES
  source                                  1..1750
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"

```



```
/clone="FLC0974"  
/tissue_type="liver"  
/dev_stage="fetus"  
379. .677  
/rpt_family="Alu"  
/rpt_type=dispersed  
complement(1163. .1396)  
/note="predicted protein of HQ2955"  
/codon_start=1  
/product="PRO2955"  
/protein_id="AAF69662.1"  
/db_xref="GI:7770253"  
/translation="MPFNCLLIGFCSSFLILLPLVCPSLVLSGNRSSWCFSEKSOPEV  
KFKCRILWLVLLIWRPAEMLVWLYVINYL"
```

ORIGIN

```
Query Match 22.1%; Score 221.6; DB 3; Length 1750;  
Best Local Similarity 80.7%; Pred. No. 2.3e-24;  
Matches 271; Conservative 0; Mismatches 64; Indels 1; Gaps 1;  
  
Qy 437 ACATTAACAGGTAAAGGCTGGGCGAGTGGCTCACACCTGTATATCCAGCACTTT 496  
Db |||||  
Qy 359 ACACGAAGGTGGGGAATAGCGCGGCGAGTGGCTCACGCTGTATATCCAGCACTTT 418  
Db |||||  
Qy 497 GGGAGGCTGAGCGAGGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCACAT 556  
Db |||||  
Qy 419 GGGAGGCGCGGAGGAGTGGCTGAGGTCAGGAGTTTGAGACCACTGACCAACAT 478  
Db |||||  
Qy 557 GCGGAACTCTGTCTCTAAACAAAATAACAAAATTAGCCTGGCATGGTGGCGGCGCT 616  
Db |||||  
Qy 479 GGTAACCCCTGTCTCTACTAATAATACAAAATTAGCCAGTGTGGTGGTGGCT 538  
Db |||||  
Qy 617 GTATCTCAGTCTGCTAGGAGCGGAGACACAAATCACTTAACCCAGAGGTGGAGG 676  
Db |||||  
Qy 539 GTATATCCAGCTACTCAGGAGGCTGAGGAGGAGAACTCACTTGAACCCGAGGAGG 598  
Db |||||  
Qy 677 TTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 735  
Db |||||  
Qy 599 TTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 658  
Db |||||  
Qy 736 TCTCAAAAGAAAAAACAAGTAAAAAGAAA 771  
Db |||||  
Qy 659 CTCAAAAAGAAAAAAGATGAGGAATATCA 694  
Db |||||
```

RESULT 4

```
CD370173 769 bp mRNA linear EST 05-AUG-2004  
LOCUS  
DEFINITION UI-H-FTI-bkb-e-05-0-UI.s1 NCI CGAP_FTI Homo sapiens CDNA clone  
UI-H-FTI-bkb-e-05-0-UI 3', mRNA sequence.  
ACCESSION CD370173  
VERSION 1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 769)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: Dr. Gary W. Hunninghake, U of I  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/cgap.html  
The following repetitive elements were found in this cDNA  
sequence: 346-636, >ALU  
Seq primer: M13 FORWARD
```

```
FEATURES  
source  
POLYA=Yes  
Location/Qualifiers  
1. .769  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="UI-H-FTI-bkb-e-05-0-UI"  
/tissue_type="Alveolar Macrophage"  
/dev_stage="Adult"  
/lab_host="DH10B (Life Technologies)"  
/clone_lib="NCI CGAP FTL"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
NCI CGAP FTL is a normalized cDNA library constructed from  
a pool of 81 RNA samples from Alveolar Macrophages  
challenged with different treatments. The mRNA samples  
were a mixture of these conditions (times refer to  
incubations following isolation by bronchoalveolar lavage)  
(some normal donor macrophages were cultured in some of  
the conditions, other donor macrophages in different  
conditions). The mRNA samples were pooled for library  
construction. Control 0 hours; control 3 hours; control 24  
hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours;  
PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella  
moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus  
moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral  
vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector  
(Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500,  
3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3  
hours; wt adenovirus + LPS 24 hours. The library was  
normalized according to Bonaldo, Lennon and Soares, Genome  
Research, 6:791-806, 1996. First strand cDNA synthesis was  
primed with an oligo-dT primer containing a Not I site.  
Double stranded cDNA was ligated to an EcoR I adaptor,  
digested with Not I, and cloned directionally into  
pT73-Pac vector. The oligonucleotide used to prime the  
synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(d)18 tail. The sequence tag for this library is  
GGCCATGCCG. The tissue was provided by Dr. Gary W.  
Hunninghake of the University of Iowa.  
TAG_TISSUE=Human Lung Alveolar Macrophage  
TAG_LIB=UI-H-FTI  
TAG_SEQ=GGCCATGCCG"
```

ORIGIN

```
Query Match 22.0%; Score 220; DB 6; Length 769;  
Best Local Similarity 80.4%; Pred. No. 5.2e-24;  
Matches 270; Conservative 0; Mismatches 65; Indels 1; Gaps 1;  
  
Qy 437 ACATTAACAGGTAAAGGCTGGGCGAGTGGCTCACACCTGTATATCCAGCACTTT 496  
Db |||||  
Qy 326 ACACGAAGGTGGGGAATAGCGCGGCGAGTGGCTCACGCTGTATATCCAGCACTTT 385  
Db |||||  
Qy 497 GGGAGGCTGAGCGAGGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCACAT 556  
Db |||||  
Qy 386 GGGAGGCGCGGAGGAGTGGATCACCTGAGGCGAGGAGTTTGAGACCACTGACCAACAT 445  
Db |||||  
Qy 557 GGCGAACTCTGTCTCTAAACAAAATAACAAAATTAGCCTGGCATGGTGGCGGCGCT 616  
Db |||||  
Qy 446 GGTAACCCCTGTCTCTACTAATAATACAAAATTAGCCAGTGTGGTGGTGGCT 505  
Db |||||  
Qy 617 GTATCTCAGTGTCTGAGGAGCGGAGACACAAATCACTTAACCCAGGAGGTGGAGG 676  
Db |||||  
Qy 506 GTATCCAGCTACTCAGGAGGCTGAGGAGGAGAACTCACTTGAACCCGAGGAGGAGG 565  
Db |||||  
Qy 677 TTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 735  
Db |||||  
Qy 566 TTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 625  
Db |||||  
Qy 736 TCTCAAAAGAAAAAACAAGTAAAAAGAAA 771  
Db |||||
```

```
Db 626 CTCAAAAAGAAAAAGATGAGAAATATCA 661

RESULT 5
BX491983
LOCUS
DEFINITION DXFzp78100513_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
VERSION DXFzp78100513 5', mRNA sequence.
KEYWORDS BX491983
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 500)
JOURNAL Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
COMMENT EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Heinrich-
Heine-University, Dueseldorf/Germany) within the cDNA sequencing
consortium of the German Genome Project. No sl sequence available.
This clone (DKFzp78100513) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
location/Qualifiers
1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFzp78100513"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI;
cDNA-collection"

ORIGIN
Query Match 21.9%; Score 219.6; DB 5; Length 500;
Best Local Similarity 82.7%; Pred. No. 7e-24; Indels 1; Gaps 1;
Matches 263; Conservative 0; Mismatches 54;

Qy 443 AACACAGGTAAAGAGGCTGGCGCAGTGGCTTCACACCTGTAAATCCAGCACTTTGGGAGG 502
Db 13 ATAAAGGAATAGAGGCTGGCGCATGGTGGCTCACGCCGTGTAATCCAGCACTTTGGGATG 72
Qy 503 CTGAGGAGGCGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCTGGCCAAATGCGGAA 562
Db 73 CCGAGGGCGGCGAGATCACCTGAGTTCAGCGCTTCGAGACTAGCTGGCCAAATGCTGAA 132
Qy 563 ACTGTGCTCTAAAAAATAACAAAATAGCTGGCATGGTGGCGGGCGCTGTAATC 622
Db 133 ACCCCATCTTTACTAAAAAATCAAAAATAGCCGGGTGGTGGTGGAGCCCTGTATC 192
Qy 623 TCAGCTGCTCAGGAGGCGGAGACACAGAAATCACTTAAACCCAGAGGTGGAGGTTCGAG 682
Db 193 CCAGCTACTTGGGAGGCTGAGAGGAGAGATCGCTTGAACGTAGAGGTGGAGGTTCGAG 252
Qy 683 TGAGCTGAGATCGTGCCACT-CATCCAACTGGGAGACAGAGTGACACTTTTCTCTCAA 741
Db 253 TGAGCTGAGATTTGTGCCACTGCTAGCTGAGTGGGTGACAGAGTGAGACTCTTGTCTCAA 312
Qy 742 AAGGAAAAAACA 759
Db 313 AAAAAA 330
```

```
BE790769 837 bp mRNA linear EST 20-SEP-2000
601581950F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3936325 5',
mRNA sequence.
BE790769
VERSION BE790769.1 GI:10211967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 837)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM780 row: c column: 14
High quality sequence stop: 723.

FEATURES
source
location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3936325"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
Query Match 21.8%; Score 218.4; DB 2; Length 837;
Best Local Similarity 70.0%; Pred. No. 8.9e-24;
Matches 308; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

Qy 332 AAGACAGTCAGGACCCCTGTTCTCACAGAGCTCATACCTAGTAGTGGTGTAGTAG 391
Db 472 AAAAAAGTGAAGCCCTTGTTATGGGCTGAATCATGCTGCCTGAATTCCTATGTTGAAG 413
Qy 392 AAATAATGCTGAGTGTCTATGTCATTTCCAGTTTTTTAGTAGCCACATTAAACAGGTA 451
Db 412 CCCTAACCCAGTAGCTCAGATGTGACTGTATTCATATAGGGCTTTAAGGGGTA 353
Qy 452 AAAAAAGCTGGCGCAGTGGCTCACACTGTGTAATCCAGCACTTTGGAGGCTCAGGCAG 511
Db 352 ATTAAGGCTGGTGCAGTGGCTCACACTATATCCCATCATCTTTGGAGGCCGAGGAG 293
Qy 512 GCAGATCACCTTTGGTCAGGAGTTGAGACTAGCTGGCCCAACATGGGAACTCTGCT 571
Db 292 GTGTATCACCTGAGGTGAGAGTTTGAGACAGGCTGACCAACATGAAGAACCTGTCT 233
Qy 572 CTAATAAAAAAATACAAAAATTTAGCCTGGCATGGTGGCGGGCGCTGTAATCTCAGCTGCT 631
Db 232 CTACTAAAAAATACAAAAATTTAGCTGTGGTGTGATACAGGCTGTAAATTCAGCCACT 173
Qy 632 CAGGAGCCGAGACACAAGAAATCACTTAACCCAGGAGGTGGAGTTGCAGTAGCTGAG 691
```

Db	433	AGTTGGAGAC	CAGCCTGGCCAA	CATGTCCTCTCT-AC	TAAAAATACAAAAA	378		
Qy	592	TAGCCTGGCA	TGTGGCGGCG	CTGTAATCTCAG	TGCTCAGAGGCG	CAGACACA	651	
Db	374	TAGCCAGGAG	TGTGGCGCAT	GCCTATATATCC	AGCTACTTTAG	GAGGCTGAGG	CAGAGA	315
Qy	652	ATCATTAAAC	CCAGGAGGTG	GGAGTTGCATG	AGCTGAGATC	GTGCCACT-CA	CTCCAA	710
Db	314	ATTACTTGA	ACCTGGGAGG	TGGAGTTGCAG	TGAGCCAA	GATTGCACCA	CTACCTCC	255
Qy	711	CTTGGGAG	CAGAGTGAC	ACTTTTGTCTC	AAAAAGAAAA	AAAAAACA	AGTAAAAA	770
Db	254	CTTGGGAG	CAAAAGTGAG	C-TCCGTCTC	AAAAA	AAAAA	AAAAA	196
Qy	771	ACAGGTGA	AGCTTAAC	TTAA	790			
Db	195	AAC	TTCCAGTTAC	CTAGAA	176			
RESULT 8								
CD518470								
LOCUS								
DEFINITION	CD518470	1019 bp	mRNA	linear	EST	06-JUN-200		
ACCESSION	AGENCOURT_14369329	NIH_MGC_181	Homo sapiens	cdna	clone			
VERSION	IMAGE:30396254	5', mRNA	sequence.					
KEYWORDS	CD518470							
SOURCE	CD518470.1	GI:31450188						
ORGANISM	Homo sapiens	(human)						
REFERENCE	Homo sapiens							
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
COMMENT	NIH-MGC http://mgs.nci.nih.gov/.							
	1 (bases 1 to 1019)							
	National Institutes of Health, Mammalian Gene Collection (MGC)							
	Unpublished (1999)							
	Contact: Daniela S. Gerhard, Ph.D.							
	Office of Cancer Genomics							
	National Cancer Institute / NIH							
	Bldg. 31 Rm10A07 Bethesda, MD 20892							
	Email: cgabs-r@mail.nih.gov							
	Tissue Procurement: Dr. Michael Brownstein							
	cdna Library Preparation: Invitrogen Corp							
	cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)							
	DNA Sequencing by: Agencourt Bioscience Corporation							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LLNL at:							
	http://image.llnl.gov							
	Plate: NDAM477 row: k column: 15							
	High quality sequence stop: 558.							
FEATURES								
source	1..1019							
	/organism="Homo sapiens"							
	/mol_type="mRNA"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:30396254"							
	/tissue_type="White Matter"							
	/dev_stage="Unknown"							
	/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"							
	/clone_lib="NIH_MGC_181"							
	/notes="vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV							
	(destroyed); Library is oligo-dT primed and directionally							
	cloned (EcoRV site is destroyed upon cloning). Average							
	insert size 1.42 kb. Library was constructed by							
	(Invitrogen). Note: this is a NIH_MGC Library."							
ORIGIN								
Query Match								
Best Local Similarity	21.6%	Score	216;	DB	6;	Length	1019;	
Matches	73.2%	Pred. No.	1.9e-23;					
Conservative	287;	Mismatches	104;	Indels	1;	Gaps	1;	
Qy	433	AGCCACATTA	AAAAACAGGTAA	AAAAAGGCTGG	CGCGAGTGGCT	CACACCTGT	TAATCC	492
Db	221	ACCCAGAGTA	AAAGAGTTGA	AGAGGCCAGG	TGCAGTGGCT	CATGCTGT	TAATCC	280

/lab host="DH10B (phage-resistant)"
/clone lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

ORIGIN

Query Match 21.5%; Score 215; DB 7; Length 633;
Best Local Similarity 83.4%; Pred. No. 3.2e-23;
Matches 256; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
Qy 454 AAAGCGTGGGCGCAGTGGCTCACACCTGTAATCCAGCACATTTGGGAGGCTGAGCGAGC 513
Db 308 ACAGCGCGGGCGCAGTGGCTCACGCCCTGTAATCCAGCATTTTGGGAGGCGGAGCAGGT 249
Qy 514 AGATCACCTTTGGTCAGGAGTTTGAGACTGACCTGGCCCAACATGGCGAAACTCTGTCTCT 573
Db 248 GGATCACCTGAGGTGAGGAGTTTGAGACCACTGGCCCAACATGGTGAACCCGCTCT 189
Qy 574 AAAAAAATACAAAAATAGCTGGCATGGTGGCGGCGCTGTATCTCAGCTGTCTCA 633
Db 188 ACTAAAAAATACAAAAATAGCTGGCATGGTGGCGTGCACCTGTATCTCAGTACTCA 129
Qy 634 GGAGCGCGAGACACAGATCACTAAACCCAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 693
Db 128 GGAGGCTGAGCGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 69
Qy 694 CGTGCCACT-CACCTCCAACTGGGAGACAGAGTGACACTTTTGTCTCAAAAGAAAAA 752
Db 68 CACACCTGCACTCCAGCCTGGGTGACAGAGTGAGACTGTCTCAAGAGGAAAAA 9
Qy 753 AAAACAA 759
Db 8 AAAAAA 2

RESULT 15

AW516097/c
LOCUS
DEFINITION
xt60f05.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:2790849 3,
similar to contains Alu repetitive element; mRNA sequence.
AW516097
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 491)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 424.

FEATURES

source
1..491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2790849"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab host="DH10B"
/clone_lib="NCI-CGAP Ut2"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.85 kb. Life Technologies catalog #:
11539-012"

ORIGIN

Query Match 21.4%; Score 214; DB 2; Length 491;
Best Local Similarity 79.4%; Pred. No. 5e-23;
Matches 278; Conservative 0; Mismatches 70; Indels 2; Gaps 2;
Qy 428 TTAGTAGCCACATTAAACACAGGTAAAGAGCTGGGCGCAGTGGCTCACACCTGTATCC 487
Db 481 TTAGTGTATTCTTTGAAAAACATGTACTTAGGCTGGGCGGTGGCTCACGCTGTATCC 422
Qy 488 CAGCACCTTTGGAGGCTGAGGCGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCCT 547
Db 421 CAGCACCTTTGGAGGCTGAGGAGTAGGTGATCAGCTGAGGTGAGGAGTTTGAGACAGCCT 362
Qy 548 GGCCAAACATGGCGAACTCTGTCTCTTAAAAAATAACAAAAATAGCCTGGCATGTGG 607
Db 361 GACCAAAAAGGTCAAACTCCGTCTCT-ACTAAAAATACAAAAATTTGGCCAGCATGGTAG 303
Qy 608 CGGCGCGCTGTATCTCAGCTGCTCAGGAGCGCAGACACAAGATCACTTAAACCCAGG 667
Db 302 CAGACGCTGTAGTTCAGCTATTTCAGGAGGATGAGACAGGAGAAATTGCTTGAACCCAGG 243
Qy 668 AGGTGGAGGTGGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 726
Db 242 AGGTGGAGGTGGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 183
Qy 727 ACACCTTTGTCTCAAAAGAAAAAACAAGTAAAAAAGAAACAGGT 776
Db 182 AGACTCTTTGTCTCCAAAAAAGAAAGAAAAAGAAAAAACAATGT 133

Search completed: July 4, 2005, 15:07:07
Job time : 2183.56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 18:09:16 ; Search time 2885.07 Seconds

(without alignments)
16811.954 Million cell updates/sec

Title: US-09-936-271C-13_COPY_8000_9000

Perfect score: 1001

Sequence: 1 aacagagaagacacatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hhg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	6	CQ788219 Sequence
2	1001	100.0	11570	6	CQ874885 Sequence
3	1001	100.0	11570	6	CQ874961 Sequence
4	1001	100.0	11570	9	AF135028 Homo sapi
5	1001	100.0	107487	9	AC011483 Homo sapi
6	1001	100.0	230000	9	AF243527 Homo sapi
7	983.6	98.3	217346	2	AC027602 Homo sapi
8	978.6	97.8	200792	2	AC130782 Pan trogl
9	289.4	28.9	165436	9	HS160H22 Human DNA
10	289.4	28.9	230141	2	AL359177 Homo sapi
11	286	28.6	129866	9	AL138703 Human DNA
12	285.4	28.5	230141	2	AL359177 Homo sapi
13	285	28.5	194197	9	AC064871 Homo sapi
14	284.2	28.4	98868	9	AL139231 Human DNA
15	283.4	28.3	143947	9	AC012462 Homo sapi
16	282.8	28.3	168686	2	AC084008 Homo sapi
17	282.6	28.2	172747	9	AL359835 Human DNA
18	282.4	28.2	26590	9	AC090672 Homo sapi
19	282.4	28.2	71506	9	AC092452 Homo sapi

C 20	282.4	28.2	162712	2	AC026096 Homo sapi
C 21	282.4	28.2	206533	9	CNS01RH1 Human chr
C 22	282.2	28.2	59257	9	AC106019 Homo sapi
C 23	282.2	28.2	65545	2	AC139371 Homo sapi
C 24	282.2	28.2	164282	2	AC015734 Homo sapi
C 25	282.2	28.2	196037	9	AC129492 Homo sapi
C 26	281.8	28.2	174654	9	AC073167 Homo sapi
C 27	281.6	28.1	166459	2	AC016092 Homo sapi
C 28	281.2	28.1	129685	9	HS01824H1 Human DNA
C 29	281.2	28.1	149445	9	AC013264 Homo sapi
C 30	281	28.1	173844	9	AC078842 Homo sapi
C 31	280.8	28.1	224573	9	AC008758 Homo sapi
C 32	280.6	28.0	152457	9	AC122693 Homo sapi
C 33	280.6	28.0	162401	2	AC024024 Homo sapi
C 34	280.6	28.0	172272	9	AC027804 Homo sapi
C 35	280.4	28.0	144683	9	AC096554 Homo sapi
C 36	280.4	28.0	146443	9	AC007488 Homo sapi
C 37	280.4	28.0	172092	2	AC013341 Homo sapi
C 38	280.4	28.0	184439	2	AP003085 Homo sapi
C 39	280.2	28.0	155217	9	AC105343 Homo sapi
C 40	280.2	28.0	161010	2	AC146128 Pan trogl
C 41	280.2	28.0	168203	2	AP001902 Homo sapi
C 42	280	28.0	186113	9	AC129097 Papio anu
C 43	280	28.0	189315	9	AC098848 Homo sapi
C 44	279.8	28.0	80658	9	AL137021 Human DNA
C 45	279.8	28.0	159934	9	AL160273 Human DNA

ALIGNMENTS

RESULT 1					
CQ788219	CQ788219	Sequence 2 from Patent WO2004021008.	11570 bp	DNA	linear PAT 24-MAR-2004
LOCUS	CQ788219				
DEFINITION	CQ788219				
ACCESSION	CQ788219				
VERSION	CQ788219.1	GI:45723068			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Diamandis,E.P.				
TITLE	Methods for detecting breast and ovarian cancer				
JOURNAL	Patent: WO 2004021008-A 2 11-MAR-2004;				
FEATURES	MOUNT SINAI HOSPITAL (CA); Yousef, George (CA)				
Location/Qualifiers					
source	1..11570				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN					
Query Match	100.0%	Score 1001;	DB 6;	Length 11570;	
Best Local Similarity	100.0%	Pred. No. 2.6e-246;			
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCAACTTT	60		
Db	8000	AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCAACTTT	8059		
QY	61	GGGAGGCGGAGCGGGTGGATCACTAGGTACAGGAGATGAGACCATCTGCTGACACGG	120		
Db	8060	GGGAGGCGGAGCGGGTGGATCACTAGGTACAGGAGATGAGACCATCTGCTGACACGG	8119		
QY	121	TGAACCCCTGTCTTCTACTAAAACACAAAAATTAGCCGGCGGTGGTGGCGGCCCTGT	180		
Db	8120	TGAACCCCTGTCTTCTACTAAAACACAAAAATTAGCCGGCGGTGGTGGCGGCCCTGT	8179		
QY	181	AGTCCCACTACTCGGAGGCGGTGAGGAGGAGATGGCTGAACCCGGAGCGGAACCTT	240		
Db	8180	AGTCCCACTACTCGGAGGCGGTGAGGAGGAGATGGCTGAACCCGGAGCGGAACCTT	8239		

[illegible]

/db_xref="GI:4589283"
/translation="MATARPFWMMVLICALITALLGVTEHVLANNNDVSCDHPNSTVPS
GSNDLGAAGEDARDSSRIINGSDMDHTOPMOAALLRNPOLYCCAVLVHPOM
LLTAHCKRVPRLVSHSYPVESQOMFQVKSI PHEGVSHPCHNSNDLMLIKLN
RRIPRTKDVRLNVSSHCPBSAGTKLVSGWGTKSPVHPKVLQCLINISVSKRCE
DAYFRQDDTNFCAKDAGKGRSDSCGDSGGPVVNCGLVSGMDYFCARPNRPGVIT
NLCKFTKWIQETIOANS"
ORIGIN

Query Match 100.0%; Score 1001; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 2.6e-246; Indels 0; Gaps 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 60
Db 8000 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 8059
Qy 61 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 120
Db 8060 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 8119
Qy 121 TGAACCCCTGCTCTACTAAAAACACAAAAAATTAGCCGGCGCTGGTGGAGCGCCTGT 180
Db 8120 TGAACCCCTGCTCTACTAAAAACACAAAAAATTAGCCGGCGCTGGTGGAGCGCCTGT 8179
Qy 181 AGTCCCGACTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240
Db 8180 AGTCCCGACTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTT 8239
Qy 241 CGAGTGAGCGGAGGTGACCACTGCACCTCGAGCTGGGCAACACAGTGAGACTCCGCT 300
Db 8240 CGAGTGAGCGGAGGTGACCACTGCACCTCGAGCTGGGCAACACAGTGAGACTCCGCT 8299
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATACATCTCAATCAAGTGGTGGCATTTAAAC 360
Db 8300 CAAAAAAGAAAAAGAAAAAGAAATACATCTCAATCAAGTGGTGGCATTTAAAC 8359
Qy 361 TATTAGGCTTCTGTAGGCAAGGTAGTATCTTGTGTTTTCCAGACCTCAAGGTGTTTT 420
Db 8360 TATTAGGCTTCTGTAGGCAAGGTAGTATCTTGTGTTTTCCAGACCTCAAGGTGTTTT 8419
Qy 421 TTGTTTGTGTTTTTATACCGGTGTGTGCTGCGGTGGGCCACACAGTGAGACTCCGCT 480
Db 8420 TTGTTTGTGTTTTTATACCGGTGTGTGCTGCGGTGGGCCACACAGTGAGACTCCGCT 8479
Qy 481 AAATAATAACACTACACAACTACTAATACCAATAGTATTAATAATATAGCATCTGCTA 540
Db 8480 AAATAATAACACTACACAACTACTAATACCAATAGTATTAATAATATAGCATCTGCTA 8539
Qy 541 ATTGCTGGACACTGTTTTAAGTGGTTTGCATGCGCTCAGCTCATTAACCTCATTTACCTGTT 600
Db 8540 ATTGCTGGACACTGTTTTAAGTGGTTTGCATGCGCTCAGCTCATTAACCTCATTTACCTGTT 8599
Qy 601 ATATTGCGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCATTAACATAAGGCTC 660
Db 8600 ATATTGCGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCATTAACATAAGGCTC 8659
Qy 661 TCAAAAGAACTCTGCAGAGATATTAATTTTAAATAATAGAGAAATTAACCCAA 720
Db 8660 TCAAAAGAACTCTGCAGAGATATTAATTTTAAATAATAGAGAGAAATTAACCCAA 8719
Qy 721 GAAAGTTGAAATTTAGAGGTACAGGAGTAAAGCTTGTGCTTTGAAACAGTGTCTGCT 780
Db 8720 GAAAGTTGAAATTTAGAGGTACAGGAGTAAAGCTTGTGCTTTGAAACAGTGTCTGCT 8779
Qy 781 ACTGGGAAAAAGGCAAGCTTTGGCTTTCTTAATAATTGATACAGGACTCTGTAAATCC 840
Db 8780 ACTGGGAAAAAGGCAAGCTTTGGCTTTCTTAATAATTGATACAGGACTCTGTAAATCC 8839
Qy 841 ATTTTGCATGCTGTAACTAGTAAGTAAGCGGGTGAATGGCACATGCCAGTAATCCC 900
Db 8840 ATTTTGCATGCTGTAACTAGTAAGTAAGCGGGTGAATGGCACATGCCAGTAATCCC 8899

Qy 901 AGCACTCTGGAGACTGAAGTGGAGAGATCACTTGAGCTCAGGAGTTCAAGACCGCTG 960
Db 8900 AGCACTCTGGAGACTGAAGTGGAGAGATCACTTGAGCTCAGGAGTTCAAGACCGCTG 8959
Qy 961 GGCACCTAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1001
Db 8960 GGCACCTAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9000
RESULT 5
AC011483/c
LOCUS AC011483 107487 bp DNA linear PRI 01-JUL-2002
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
AC011483
AC011483.7 GI:21637461
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107487)
DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
1. 107487
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C22"
64998..65494
/note="NOTE: Shatter libraries failed to resolve
dinucleotide repeat. Unsure number of repeat copies
64998-65494. Forced join 65015."
misc_feature
ORIGIN
Query Match 100.0%; Score 1001; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 3.2e-246; Indels 0; Gaps 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 60
Db 52220 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 52161
Qy 61 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 120
Db 52160 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 52101
Qy 121 TGAACCCCTGCTCTACTAAAAACACAAAAAATTAGCCGGCGCTGGTGGAGCGCCTGT 180
Db 52100 TGAACCCCTGCTCTACTAAAAACACAAAAAATTAGCCGGCGCTGGTGGAGCGCCTGT 52041
Qy 181 AGTCCCGACTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240


```
CDS      /product="glandular kallikrein"
join(61139..61184,62391..62550,64142..64428,64542..64678,
66074..66229)
/gene="KLK2"
/note="serine protease"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MWDLVLSIALSVGCTGAVPLIOSRIVIGMECEKHSOPMOVAVYS
HGWAHCGVLHPQWLVITAACHLKNQSVWLGRHNLPEPDTGQRVPVSHSFPHPLYN
MSLLKHSQSLRPDSDHLLMLLSEPAKITDVVVLGLPTQBPALGTTCYASGWSGI
EPEFLRLSLQCVSLHLLNDMCARAYSEKVTFFMLCAGLTGKGDKTCGSDSGGPLV
CNGVLOGITSGPCCALPRKPAVYTKVHYRWIKDTIAANP"
complement(join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..98405))
/product="serine protease"
complement(join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPPWGLFYLILGVAGSLVSGSCSOIINGDCSPHSQPW
QAALVMEELFCGVLVHPQWLVSAACHFQNSYITIGLHSLDADQEPGSGQVSEASLS
VRPEYNRLIANDMLIKLDESVESEDITRSISIAQCPTAGNSCLVSGWGLIANGR
MPTVLCQVNVSWVEECSKLYDPLYPHPSMFCAGGHDQKDCNGSDSGGPLICNGYLIQ
GLVSFGAPCGQGVPGVYTNLCKFTWIEKTVQAS"
complement(<131301..>140303)
/gene="KLK5"
/note="synonym: SCTE"
complement(join(<131301..131456,136310..136443,
136529..136785,137525..137690,140255..>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement(join(131301..131456,136310..136443,
136529..136785,137525..137690,140255..>140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWVLICALLITALLGVTDARSDDSSRIINGSDDMHTQPWQAAL
LLRNQLYCGAVLHPQWLVTAACHKKVFRVLGHYSLSFVBSGQOMQGVKSIHP
PGYSHPGHNDMLIKLNRIRPTKDVPIVNSHCPISAGTKCLVSGWGTCKSPQVHF
PKVQLCNISVLSKREDAYPROIDTMEFCAGBKAGRDCSQGSGFPVVCNGSLQGL
VSWGDYPCARPNRFGVYTNLCKFTWIEKTVQAS"
complement(join(<146834..146986,149628..149764,
151186..151433,155052..155208,155948..>155987))
/product="protease M"
complement(join(146834..146986,149628..149764,
151186..151433,155052..155208,155948..>155987))
/note="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKKLMMVLSLIAAAWAEQNLVHGPGCDKTSHPYQAALYTSGH
LLCGVLHPUNWVITAACHKPNLQVFLGKHLNRQRESSQSSQSVRAVHPDYDAAS
HDQRLMLRLARPAKLISLIQPLERDCSANTTSCHILGWKGTADGFPDTIQDAYI
HLVRECEHAYPQGITQNMCLCAGDEKYGKDSQCQSDSGGPLVCGDHLRLVSWGNIPC
GSKBPGVYTNVCRVYTNWIKTIQAK"
complement(<165420..>170283)
/gene="KLK7"
/note="synonym: SCTE"
complement(join(<165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement(join(165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283))
/gene="KLK7"
/note="serine protease"
mRNA     /codon_start=1
/translation="MARSLLLPLQIILLLSLALETAGEAQQDKIIDGAPCARGSHPNQ
VALLSGNQLHCGGVNLNRWLVTAACHKNEYTVHLGSDTLGDRRAQRIKASKSFRHP
GYSTQTHVNDLMLVNLKNSQARLSMVKVLRPSRCEPPGTCTVSGMGTTSPDVTFF
SDLMCVDKLISPDQCTKVYKDLLENSMLCAGIPDSKKNACNGSDSGGPLVCRGTLQGL
VSMGTTPCGPNPDPGVYTVCKFTKWNIDTMKKHR"
complement(join(<183943..184098,185635..185768,
187865..188127,188293..188452,188967..>189036))
/product="neuropsin"
complement(join(183943..184098,185635..185768,
187865..188127,188293..188452,188967..>189036))
/note="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRPRPRAAKTWFFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQ
PWQAALFQGOQLLCGGVLVGNWLVTAACHKPKYTVRLGDHSLQNDGPEQEIPIVVQ
SIHPCYNSDVEDHNDHMLLQLRDQASLGSVKVPLSLADHCTQPCQKCTVSGWGTV
TSRENPPDTLNCAEVKIIFQKKCEDAYPCQITDGMVCSAGSSKAGATCQGDSSGGLVC
DGLQGITSGSDPCGRSDRPGVYTNICRYLDWIKKILGSKG"
complement(join(<190980..191129,191573..191709,
194324..194589,197048..197204,197370..>197412))
/product="kallikrein-like 3"
complement(join(190980..191129,191573..191709,
194324..194589,197048..197204,197370..>197412))
/note="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translation="MKLLGALLSLLAGHWADTRAIABECPNSQPWQAGLFLHLT
RLFCGATLSDRWLLTAACHKPKYLVWLGRHLLWKEGPEQLFRVTFDFPFGFNKD
LSANDHNDIMLIRLPQARLSPAVPLNISQTSFPMQCLISQVSGAVSSPKALFV
TLQCANISILENKLCHWAYPGHISDSMLCAGLMEGGRGSCQDSSGGLVLCVNGTLAGVV
SGNAEPCSRPRRPAVYTVCHYLDWIQIMEN"
```

```
Query Match 100.0%; Score 1001; DB 9; Length 230000;
Best Local Similarity 100.0%; Pred. No. 3,4e-246;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACAGAGAAAGACATCTCGCGCGCGGTGGCTCACACCTGTATATCCCAACACTTT 60
Db 134548 AACAGAGAAAGACATCTCGCGCGCGGTGGCTCACACCTGTATATCCCAACACTTT 134489
Qy 61 GGGAGGCGGAGCGGGTGGATCACTAGGTGAGGAGATGGAGACCATCTGGCTGACACGG 120
Db 134488 GGGAGGCGGAGCGGGTGGATCACTAGGTGAGGAGATGGAGACCATCTGGCTGACACGG 134429
Qy 121 TGAACCCCTGTCTTACTATAAACAACAAAAATTAGCCGGCGGTGGTGGAGGCGCCTGT 180
Db 134428 TGAACCCCTGTCTTACTATAAACAACAAAAATTAGCCGGCGGTGGTGGAGGCGCCTGT 134369
Qy 181 AGTCCCAGCTACTCGGAGGCTGAGGCGAGGAGATGCCTGAAACCGGAGCGCGAACTT 240
Db 134368 AGTCCCAGCTACTCGGAGGCTGAGGCGAGGAGATGCCTGAAACCGGAGCGCGAACTT 134309
Qy 241 GCAGTGAAGCGAGGTTGCACCATCTGCACCTCAGCCTGGGCAACACAGTGAAGACTCCGCT 300
Db 134308 GCAGTGAAGCGAGGTTGCACCATCTGCACCTCAGCCTGGGCAACACAGTGAAGACTCCGCT 134249
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGGTGGCAATTAAAC 360
Db 134248 CAAAAAAGAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGGTGGCAATTAAAC 134189
Qy 361 TATTAGCCTTTCTGTAGGCAAGGTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTTT 420
Db 134188 TATTAGCCTTTCTGTAGGCAAGGTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTTT 134129
Qy 421 TTGTTGTTTTTTTCATACCCGCTGTGTGTGGTGGGCCACTAAAAAGCTACAAAGCAAG 480
```


* 17511 17610: gap of 100 bp
* 24602: contig of 6992 bp in length
* 24703 24702: gap of 100 bp
* 35434: contig of 10732 bp in length
* 35435 35434: gap of 100 bp
* 124474: contig of 88940 bp in length
* 124575 124574: gap of 100 bp
* 134664: contig of 10090 bp in length
* 134765 134764: gap of 100 bp
* 162343 162343: contig of 27579 bp in length
* 162344 162344: gap of 100 bp
* 208917 208917: contig of 46474 bp in length
* 208918 209017: gap of 100 bp
* 209018 217346: contig of 8329 bp in length.

FEATURES

source

1. .217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. .8149
/note="assembly_fragment"
clone_end:SP6
vector_side:left

misc_feature

8250. .9592
/note="assembly_fragment"

misc_feature

9693. .10733
/note="assembly_fragment"

misc_feature

10834. .13519
/note="assembly_fragment"

misc_feature

13620. .17510
/note="assembly_fragment"

misc_feature

17611. .24602
/note="assembly_fragment"

misc_feature

24703. .35434
/note="assembly_fragment"

misc_feature

35535. .124474
/note="assembly_fragment"

misc_feature

124575. .134664
/note="assembly_fragment"

misc_feature

134765. .162343
/note="assembly_fragment"

misc_feature

162444. .208917
/note="assembly_fragment"

misc_feature

209018. .217346
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN

Query Match 98.3%; Score 983.6; DB 2; Length 217346;
Best Local Similarity 99.5%; Pred. No.1e-241;
Matches 997; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
Qy 1 AACGAGAAAGCAGCATCTCGGCGGGCTGAGGCTCACACCTGTATATCCCAACACTTT 60
Db 196413 AACGAGAAAGCAGCATCTCGGCGGGCTGAGGCTCACACCTGTATATCCCAACACTTT 196354
Qy 61 GGGAGGCGGCGGGTGGATCACTAGGTCAGGAGTGGAGACCATCTGGCTGCACCGG 120
Db 196353 GGGAGGCGGCGGGTGGATCACTAGGTCAGGAGTGGAGACCATCTGGCTGCACCGG 196294
Qy 121 TGAACCCCTGCTCTACTAAAAACAAAAAATAGCCGGCGGTGGTGGCAGGCGCCCTGT 180
Db 196293 TGAACCCCTGCTCTACTAAAAACAAAAAATAGCCGGCGGTGGTGGCAGGCGCCCTGT 196234
Qy 181 AGTCCACGACTCTCGGAGGCTGAGGCAAGGAGATGGCCCTGAAACCCGGAGCGCGAACTT 240
Db 196233 AGTCCACGACTCTCGGAGGCTGAGGCAAGGAGATGGCCCTGAAACCCGGAGCGCGAACTT 196174

Qy 241 GCAGTGAGCGAGGTTGCACCACTGCAGCTCGGGGCAACACAGTGAGACTCCGCTCT 300
Db 196173 GCAGTGAGCGAGGTTGCACCACTGCAGCTCGGGGCAACACAGTGAGACTCCGCTCT 196114
Qy 301 CAAAAAATAA-AAAAGAAAAAGAAATCAATCTCATTCATTCAGTGGTGGCATTTAAAA 359
Db 196113 CAAAAAATAA-AAAAGAAAAAGAAATCAATCTCATTCAGTGGTGGCATTTAAAA 196054
Qy 360 CTATTTAGCCTTTCTGTAGGCAAGGTTAGTATCTTGTTTTCCAGACCTCAAGTGTTTT 419
Db 196053 CTATTTAGCCTTTCTGTAGGCAAGGTTAGTATCTTGTTTTCCAGACCTCAAGTGTTTT 195994
Qy 420 TTTGTTTGTTTTTCATACCGGTGTGTGTCTGGGTGGGCCCACTAAAGACTCAAGCAA 479
Db 195993 TTTGTTTGTTTTTCATACCGGTGTGTGTCTGGGTGGGCCCACTAAAGACTCAAGCAA 195934
Qy 480 GAAATATAACAACTACCAATACCAATACCAATAGTATAAAAAATATAGACTCTGGCT 539
Db 195933 GAAATATAACAACTACCAATACCAATACCAATAGTATAAAAAATATAGACTCTGGCT 195874
Qy 540 AATTGCTGGACACTGTTTAAAGTGGTTGCATGCCTCAGCTCATTAATCACTTACCTGT 599
Db 195873 AATTGCTGGACACTGTTTAAAGTGGTTGCATGCCTCAGCTCATTAATCACTTACCTGT 195814
Qy 600 TATTATTGGCCCTATTTTACAAACAAAGGAGCCCAAGGCTCAGAGCAGTTAACTAACAGCCT 659
Db 195813 TATTATTGGCCCTATTTTACAAACAAAGGAGCCCAAGGCTCAGAGCAGTTAACTAACAGCCT 195754
Qy 660 CTCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATATAGAGAAAAATTAACACCA 719
Db 195753 CTCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATATAGAGAAAAATTAACACCA 195694
Qy 720 AGAAGTTGAAATTTAGAGGTACAGCAGCTTAAGCTTGTTCCTTTTGAACAGAGTCTGC 779
Db 195693 AGAAGTTGAAATTTAGAGGTACAGCAGCTTAAGCTTGTTCCTTTTGAACAGAGTCTGC 195634
Qy 780 TACTGGGAAAAAGCAAGTCTTGGCTTTCCTAATAATGATACCAAGACTCTGTAATTCA 839
Db 195633 TACTGGGAAAAAGCAAGTCTTGGCTTTCCTAATAATGATACCAAGACTCTGTAATTCA 195574
Qy 840 TATTTGTCATGATGTAAAGTAAGAAATGAAGCCGGGTGCAATGGCAATGCCAATATCC 899
Db 195573 TATTTGTCATGATGTAAAGTAAGAAATGAAGCCGGGTGCAATGGCAATGCCAATATCC 195514
Qy 900 CAGCACTCTGGAGACTGGAAGTGGAGAGTCACTTGGCTCAGGAGTTCAGAGCAGCCT 959
Db 195513 CAGCACTCTGGAGACTGGAAGTGGAGAGTCACTTGGCTCAGGAGTTCAGAGCAGCCT 195454
Qy 960 GGGCACTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1001
Db 195453 GGGCACTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 195412

RESULT 8
AC130782/c 200792 bp DNA linear HTG 22-NOV-2002
LOCUS
DEFINITION Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
ordered pieces.
AC130782
AC130782.2 GI:25167101
VERSION HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 200792)
AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blackley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Cariga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghghni, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Masello, C., Maskeri, B., McDowell, J.,
Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,

Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
 Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 200792)
 Direct Submission
 Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 200792)
 Direct Submission
 Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Nov 22, 2002 this sequence version replaced gi:22218452.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: dhz
 Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 196656 bases at least Q40
 Consensus quality: 197883 bases at least Q30
 Consensus quality: 198879 bases at least Q20
 Insert size: 215000; agarose-fp
 Quality coverage: 9.03x in Q20 bases; agarose-fp
 Quality coverage: 9.73x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced
 by the finished sequence as soon as it is available and the accession number will be preserved.

1 28306: contig of 28306 bp in length
 28307 28406: gap of unknown length
 28407 37856: contig of 9450 bp in length
 37857 73522: contig of unknown length
 73523 73622: contig of 35566 bp in length
 73623 83567: gap of unknown length
 83568 83667: contig of 9945 bp in length
 83668 88817: gap of unknown length
 88818 88917: contig of 5150 bp in length
 88918 125611: contig of 36694 bp in length
 125612 125711: gap of unknown length
 125712 159879: contig of 34168 bp in length
 159880 159979: gap of unknown length
 159980 174698: contig of 14719 bp in length

174699 174798: gap of unknown length
 174799 186382: contig of 11584 bp in length
 186383 186483: gap of unknown length
 186483 193344: contig of 6862 bp in length
 193345 193445: gap of unknown length
 193445 193633: contig of 5919 bp in length
 193634 194633: gap of unknown length
 194634 200792: contig of 1329 bp in length.
 FEATURES
 source
 1..200792
 /organism="pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="CH251-355A20"
 /clone_lib="CH251"
 1..28306
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 28407..37856
 /note="assembly_fragment"
 37957..73522
 /note="assembly_fragment"
 73623..83567
 /note="assembly_fragment"
 83668..88817
 /note="assembly_fragment"
 88918..125611
 /note="assembly_fragment"
 125712..159879
 /note="assembly_fragment"
 159980..174698
 /note="assembly_fragment"
 174799..186382
 /note="assembly_fragment"
 186483..193344
 /note="assembly_fragment"
 193445..199363
 /note="assembly_fragment"
 199464..200792
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right

ORIGIN

Query Match 97.8%; Score 978.6; DB 2; Length 200792;
 Best Local Similarity 98.6%; Pred. No. 1.9e-240;
 Matches 987; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 1 AACAGAGAAAGCACATCTCGGCCGGCGTGGTGCCTCACACCTGTATATCCCAACACTTT 60
 Db 71159 AACAGAGAAAGCACATCTCGGCCGGCGTGGTGCCTCACACCTGTATATCCCAACACTTT 71100
 QY 61 GGGAGGCCGAGCGGGTGGATCCTAGTCTAGGAGATGAGACCATCTCTGCTGCACGG 120
 Db 71099 GGGAGGCCGAGCGGGTGGATCCTAGTCTAGGAGATGAGACCATCTCTGCTGCACGG 71040
 QY 121 TGAACACCTGCTCTCTACTTAAACACACAAAAAATAGCCGGCGCTGGTGGCAGGCGCCCTGT 180
 Db 71039 TGAACACCTGCTCTCTACTTAAACACACAAAAAATAGCCGGCGCTGGTGGCAGGCGCCCTGT 70980
 QY 181 AGTCCCGAGTCTACTCGGAGGCTGAGCGGAGGAGATGGCTTGAACCCGGGAGCGGAACTT 240
 Db 70979 AGTCCCGAGTCTACTTGGGAGGCTGAGCGGAGGAGATGGCTTGAACCCGGGAGCGGAGCTT 70920
 QY 241 GCAGTGAGCCGAGTTGCAACCACTGCCTCAGCTGGGCAACACAGTGAAGTCCGCTCT 300
 Db 70919 GCAGTGAGCCGAGTTGCAACCACTGCCTCAGCTGGGCAACACAGTGAAGTCCGCTCT 70860
 QY 301 CAAAAAAGAAAAAGAAAAAGAAAAAGAAATCATCTCTCAATGTTGGGCAATTTAAAC 360
 Db 70859 CAAAAAAGAAAAAGAAAAAGAAAAAGAAATCATCTCTCAATGTTGGGCAATTTAAAC 70800

```
QY 361 TATTAGCCTTCTGTAGGCAAGGTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTT 420
Db 70799 TATTAGCCTTCTGTGGGCAAGGTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTT 70740
QY 421 TTGTTTGTGTTTTTCATACCGGTGTGTGCTGCGGTGTGGCCACTAAAAGCTTACAAGCAAG 480
Db 70739 TTGTTTGTGTTTTTCATACCGGTGTGTGCTGCGGTGTGGCCACTAAAAGCTTACAAGCAAG 70680
QY 481 AAATAATAACAACCTACCAACAATCTAATPACCAATAGTATAAAAAATAATAGCATCTGGCTA 540
Db 70679 AAATAATAACAACCTACCAACAATCTAATPACCAATAGTATAAAAAATAATAGCATCTGGCTA 70620
QY 541 ATTGCTGACACTGTTTAAAGTGTGTCATGCTCCTCAGCTCATTAATCATATTAACCTGTT 600
Db 70619 ATTGCTGACACTGTTTAAAGTGTGTCATGCTCCTCAGCTCATTAATCATATTAACCTGTT 70560
QY 601 ATTATTGCGCCCTATTATCAACCAAGGAGCAAGGCTCAGAGCAGTAACTTAACAGCCTC 660
Db 70559 ATTATTGCGCCCTATTATCAACCAAGGAGCAAGGCTCAGAGCAGTAACTTAACAGCCTC 70500
QY 661 TCAAAAAGAAACTCTGCAGAGATATTAAATTTTAAAAAATAATGAGAGAAATTAACCCAA 720
Db 70499 TCAAAAAGAAACTCTGCAGAGATATTAAATTTTAAAAAATAATGAGAGAAATTAACCCAA 70440
QY 721 GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGCTTTGAAACAGTGTCTGCT 780
Db 70439 GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGCTTTGAAACAGTGTCTGCT 70380
QY 781 ACTGGGAAAAGGCAAGCTTGGCTTCTTAATTAATGATACCAAGGCTCTGTAATTCAT 840
Db 70379 ACTGGGAAAAGGCAAGCTTGGCTTCTTAATTAATGATACCAAGGCTCTGTAATTCAT 70320
QY 841 ATTTTGCATGATGTAACTAAGATGAAGCGGCTGCAATGGCAGTCCAGTAAATCCC 900
Db 70319 ATTTTGCATGATGTAACTAAGATGAAGCGGCTGCAATGGCAGTCCAGTAAATCCC 70260
QY 901 AGCACTCTGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGCAGCCTG 960
Db 70259 AGCACTCTGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGCAGCCTG 70200
QY 961 GGCACCTAAAATTTAAAAAATAAATACTAATTTGTTTT 1001
Db 70199 GGCACCTAAAATTTAAAAAATAAATACTAATTTGTTTT 70159

RESULT 9
HSA160H22/c HSA160H22 165436 bp DNA linear PRI 01-DEC-2000
LOCUS Human DNA sequence from clone RP11-160H22 on chromosome
DEFINITION lq23.2-24.3, complete sequence.
ACCESSION AL121983
VERSION AL121983.13 GI:10443470
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165436)
Bagguley, C.
Direct Submission
Submitted (30-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Oct 1, 2000 this sequence version replaced gi:9998793.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi, EMBL; Swi, SWISSPROT; Tri, TREMBL; Wp, WORMPEP; Information
```

on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
RP11-160H22 is from the library RPCI-11.1 constructed at the
Roswell Park Cancer Institute by the group of Pieter de Jong. For
details see <http://bacpac.med.buffalo.edu/>
VECTOR: pBACe3.6
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
IMPORTANT: This sequence is not the entire insert of clone
RP11-160H22. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-160H22 is at 165436 in this
sequence. The true left end of clone RP11-413B16 is at 81534 in
this sequence. The true right end of clone RP4-681J21 is at 129090
in this sequence. The true left end of clone RP5-1198E17 is at 100
in this sequence.

FEATURES

```
source
1.165436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="q23.2-24.3"
/clone="RP11-160H22"
/clone_lib="RPCI-11.1"
1..152
/notes="AluSx repeat: matches 7..157 of consensus"
repeat_region
163..1009
/notes="L1MC3 repeat: matches 6619..7483 of consensus"
repeat_region
1110..1363
/notes="L1MC3 repeat: matches 6322..6604 of consensus"
repeat_region
1417..2313
/notes="L1MC2 repeat: matches 5179..6113 of consensus"
repeat_region
2314..2654
/notes="AluJb repeat: matches 2..310 of consensus"
repeat_region
2665..2776
/notes="AluJb repeat: matches 8..133 of consensus"
repeat_region
2807..2985
/notes="L1MC2 repeat: matches 4984..5164 of consensus"
repeat_region
3006..3320
/notes="AluJb8 repeat: matches 1..315 of consensus"
repeat_region
3341..3380
/notes="L1M4 repeat: matches 4443..4479 of consensus"
repeat_region
3381..3673
/notes="AluSx repeat: matches 1..296 of consensus"
repeat_region
3674..3779
/notes="L1M4 repeat: matches 4325..4443 of consensus"
repeat_region
3870..4196
/notes="AluJb8 repeat: matches 1..304 of consensus"
repeat_region
4243..4533
/notes="AluSx repeat: matches 1..300 of consensus"
repeat_region
4589..5103
/notes="L1M4 repeat: matches 3804..4328 of consensus"
repeat_region
5181..5263
/notes="L1MB4 repeat: matches 6095..6183 of consensus"
repeat_region
5445..5742
/notes="AluSx repeat: matches 1..298 of consensus"
repeat_region
5985..5994
/notes="15 copies 2 mer ac 100% conserved"
repeat_region
5995..6018
/notes="12 copies 2 mer ag 95% conserved"
repeat_region
6708..6987
/notes="AluSp repeat: matches 1..297 of consensus"
repeat_region
7033..7339
/notes="AluSx repeat: matches 1..307 of consensus"
repeat_region
7405..7707
```

repeat_region /note="Alusx repeat: matches 1. .303 of consensus"
7812. .8120
repeat_region /note="AluJo repeat: matches 1. .311 of consensus"
8547. .8857
repeat_region /note="Alusx repeat: matches 1. .311 of consensus"
9267. .9520
repeat_region /note="Alusq repeat: matches 1. .254 of consensus"
10244. .10539
repeat_region /note="AluJo repeat: matches 2. .303 of consensus"
complement(10251. .10663)
misc_feature /note="match: GSS: Em:AQ056360"
11287. .11622
misc_feature /note="match: GSS: Em:AQ087768"
12105. .12400
repeat_region /note="AluJo repeat: matches 1. .294 of consensus"
12406. .12701
repeat_region /note="Alusp repeat: matches 5. .311 of consensus"
complement(13018. .13405)
misc_feature /note="match: GSS: Em:AQ470631"
13041. .13258
repeat_region /note="WIR repeat: matches 11. .259 of consensus"
13309. .13442
repeat_region /note="L1MB8 repeat: matches 6034. .6171 of consensus"
13685. .13816
repeat_region /note="WIR repeat: matches 47. .177 of consensus"
14739. .15025
repeat_region /note="Alusx repeat: matches 1. .303 of consensus"
16161. .16401
repeat_region /note="L2 repeat: matches 1594. .1811 of consensus"
16402. .16677
repeat_region /note="Alusq repeat: matches 2. .278 of consensus"
16582. .16914
misc_feature /note="match: STS: Em:HS278YE5"
16678. .16885
repeat_region /note="L2 repeat: matches 1811. .2085 of consensus"
16958. .17254
repeat_region /note="AluY repeat: matches 2. .297 of consensus"
17394. .17708
repeat_region /note="AluJb repeat: matches 1. .307 of consensus"
17819. .18069
repeat_region /note="Alusq repeat: matches 1. .239 of consensus"
18343. .18456
repeat_region /note="WLR1H repeat: matches 96. .211 of consensus"
18522. .18710
repeat_region /note="WLR1H repeat: matches 347. .535 of consensus"
19219. .19529
repeat_region /note="Alusq repeat: matches 1. .311 of consensus"
19681. .19751
repeat_region /note="WIR repeat: matches 124. .197 of consensus"
20184. .20402
repeat_region /note="L1MB3A repeat: matches 5926. .6152 of consensus"
20403. .20698
repeat_region /note="AluY repeat: matches 1. .295 of consensus"
20699. .20754
repeat_region /note="L1MB3A repeat: matches 5872. .5926 of consensus"
20757. .21052
repeat_region /note="AluJo repeat: matches 1. .302 of consensus"
21990. .22293
repeat_region /note="Alusx repeat: matches 1. .304 of consensus"
22631. .22699
repeat_region /note="L2 repeat: matches 2471. .2533 of consensus"
22852. .23141
repeat_region /note="Alusx repeat: matches 1. .291 of consensus"
23167. .23290
repeat_region /note="WIR repeat: matches 112. .249 of consensus"
23452. .23493
repeat_region /note="L1MA2 repeat: matches 6267. .6308 of consensus"
23575. .23682
repeat_region /note="L2 repeat: matches 2630. .2748 of consensus"
23707. .23969
repeat_region /note="AluJb repeat: matches 40. .301 of consensus"
25226. .25536
repeat_region /note="AluJb repeat: matches 1. .293 of consensus"

repeat_region 25797. .26107
/note="Alusx repeat: matches 1. .309 of consensus"
26292. .26357
repeat_region /note="33 copies 2 mer aa 68% conserved"
26594. .26706
repeat_region /note="MER3 repeat: matches 33. .146 of consensus"
26750. .27027
repeat_region /note="AlusC repeat: matches 1. .285 of consensus"
27028. .27088
repeat_region /note="L1MC2 repeat: matches 5788. .5852 of consensus"
27092. .27182
repeat_region /note="AluJ/FLAM repeat: matches 2. .91 of consensus"
27217. .27673
repeat_region /note="L1MC2 repeat: matches 5847. .6321 of consensus"
27674. .27734
repeat_region /note="MER3 repeat: matches 147. .209 of consensus"
27695. .27745
repeat_region /note="MER3 repeat: matches 41. .89 of consensus"
28664. .28959
repeat_region /note="Alusg1 repeat: matches 1. .292 of consensus"
28986. .29101
repeat_region /note="Alusq/x repeat: matches 9. .124 of consensus"
29214. .29504
repeat_region /note="Alusx repeat: matches 1. .291 of consensus"
29721. .29852
repeat_region /note="FLAM C repeat: matches 2. .133 of consensus"
30810. .31027
repeat_region /note="L1MB5 repeat: matches 5944. .6176 of consensus"

Query Match 28.9%; Score 289.4; DB 9; Length 165436;
Best Local Similarity 88.5%; Pred. No. 2.4e-63;
Matches 314; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGGTGGTGGCTACAGCTGATGGAGACCATCTGGCTGACACG 120
Db 162103 AAAAAAAAAAAAAAAAAAGCGCGGGTGGCTACAGCTGATGGAGACCATCTGGCTGACACG 161984

Qy 61 GGGAGCGGAGGGTGGATCAGTACAGGATGGAGACCATCTGGCTGACACG 120
Db 162043 GGGAGCGGAGGGTGGATCAGTACAGGATGGAGACCATCTGGCTGACACG 161984

Qy 121 TGAACCTCTCTCTACTAAAAACACAAAAATAGCGGGCGTGGTGGAGGCGGCTGT 180
Db 161983 TGAACCTCTCTCTACTAAAAACACAAAAATAGCGGGCGTGGTGGAGGCGGCTGT 161924

Qy 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGCCTGAACCCGGAGGCGGAATT 240
Db 161923 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGCCTGAACCCGGAGGCGGAGCTT 161864

Qy 241 GCAGTGAGCGAGTTGCACCATCGACTGAGCTGGGCAACACAGTGAGACTCCGCT 300
Db 161863 GCAGTGAGCGAGATTGCGCCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCT 161804

Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATCACTCTCATTCAGTGGTGGCAATT 355
Db 161803 CAAAAAAGAAAAAGAAAAAGAAAAATTCCTCCATTCATAGTTGCCTAT 161749

RESULT 10
AL359177
LOCUS AL359177 230141 bp DNA linear HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-413B16, 16 unordered pieces.
ACCESSION AL359177
VERSION AL359177.3 GI:12331067
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Burton, J.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 22, 2001 this sequence version replaced gi:9797929.

COMMENT

----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA413B16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: pGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 127350 bases at least Q40
Consensus quality: 132189 bases at least Q30
Consensus quality: 134528 bases at least Q20
Insert size: 228641; sum-of-contigs
Insert size: 152695; 8.1% error; agarose-fp
Quality coverage: 2.04x in Q20 bases; sum-of-contigs Quality coverage: 3.11x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 3155: contig of 3155 bp in length
* 3156 3255: gap of 100 bp
* 3256 10055: contig of 6800 bp in length
* 10056 10155: gap of 100 bp
* 10156 16833: contig of 6678 bp in length
* 16834 16933: gap of 100 bp
* 16934 36991: contig of 20058 bp in length
* 36992 37091: gap of 100 bp
* 37092 46051: contig of 8960 bp in length
* 46052 46151: gap of 100 bp
* 46152 58091: contig of 11940 bp in length
* 58092 58191: gap of 100 bp
* 58192 64171: contig of 5980 bp in length
* 64172 70251: contig of 5980 bp in length
* 70252 70351: gap of 100 bp
* 70352 73351: contig of 3000 bp in length
* 73352 73451: gap of 100 bp
* 73452 79431: contig of 5980 bp in length
* 79432 79531: gap of 100 bp
* 79532 85511: contig of 5980 bp in length
* 85512 85611: gap of 100 bp
* 85612 100531: contig of 14920 bp in length
* 100532 100631: gap of 100 bp
* 100632 106611: contig of 5980 bp in length
* 106612 106711: gap of 100 bp
* 106712 12459: contig of 18748 bp in length
* 12460 12559: gap of 100 bp
* 12560 134519: contig of 8960 bp in length
* 134520 134619: gap of 100 bp
* 134620 230141: contig of 95522 bp in length.

FEATURES
source

1..230141
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-413B16"
/clone_lib="RPC1-11.2"
1..3155
/note="assembly_fragment:00110"
3256..10055
/note="assembly_fragment:00856"

misc_feature

misc_feature

misc_feature 10156..16833
/note="assembly_fragment:01004"
misc_feature 16934..36991
/note="assembly_fragment:01205"
misc_feature 37092..46051
/note="assembly_fragment:01284"
misc_feature 46152..58091
/note="assembly_fragment:01287"
misc_feature 58192..64171
/note="assembly_fragment:01291"
misc_feature 64272..70251
/note="assembly_fragment:01293"
misc_feature 70352..73351
/note="assembly_fragment:01295"
misc_feature 73452..79431
/note="assembly_fragment:01300"
misc_feature 79532..85511
/note="assembly_fragment:01302"
misc_feature 85612..100531
/note="assembly_fragment:01305"
misc_feature 100632..106611
/note="assembly_fragment:01306"
misc_feature 106712..125459
/note="assembly_fragment:01310"
misc_feature 125560..134519
/note="assembly_fragment:01312"
misc_feature 134620..230141
/note="assembly_fragment:01255
clone_end:17
vector_side:right"

ORIGIN

Query Match 28.9%; Score 289.4; DB 2; Length 230141;
Best Local Similarity 88.5%; Pred. No. 2:5e-63;
Matches 314; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
Qy 1 AACAGAGAAAGCACATCTCGGCGGCGGTGGTGGCTCACACCTGTATATCCACACTTT 60
Db 149504 AAAAAAAAAAAAAAAAAAGGCGGCGGTGGCTCACGCTGTATATCCACACTTT 149563
Qy 61 GGAGGCGGAGGCGGTGGATCACTAGGTGAGGATGAGAGACCATCTGGCTGACACGG 120
Db 149564 GGAGGCGGAGGCGGTGGATCACTAGGTGAGGATGAGAGACCATCTGGCTAACACGG 149623
Qy 121 TGAACCCCTCTCTACTTAAACACAAAAAATAGCGGCGGTGGTGGCGGCGCTGT 180
Db 149624 TGAACCCCTCTCTACTTAAACATACAAAAATAGCGGCGGTGGTGGCGGCGCTGT 149683
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGCGCTGAACCCGGGAGCGGAACCT 240
Db 149684 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGCGCTGAACCCGGGAGCGGAGCTT 149743
Qy 241 GCAGTGAGCGGAGTTGCACATGCACTCAGCTCGGCAACACAGTGAGACTCCGCTCT 300
Db 149744 GCAGTGAGCGGAGTTGCGCACTCAGCTCGGCACTCAGCTCGGCAACAGTGAGACTCCGCTCT 149803
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATACATCTCATTCAGTGGTGGCATTT 355
Db 149804 CAAAAAAGAAAAAGAAAAAGAAATTTTCCCATTTCTATAGGTGGCTAT 149858

RESULT 11

ALL138703/c ALL138703 129866 bp DNA linear PRI 21-APR-2001
LOCUS Human DNA sequence from clone RP11-413N19 on chromosome 13,
DEFINITION complete sequence.
ACCESSION AL138703
VERSION AL138703.10 GI:13785012
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 129866)
Tracey A.
Direct Submission
Submitted (21-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 24, 2001 this sequence version replaced gi:13374974.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr13>
RP11-413N19 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone
RP11-413N19 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-413N19 is at 129866 in this
sequence. The true right end of clone RP11-215B13 is at 100 in this
sequence.

FEATURES
source

Location/Qualifiers
1..129866

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-413N19"
/clone_lib="RPCI-11.2"
669..768

repeat_region
/note="Charliel repeat: matches 1455..1554 of consensus"
878..1291
repeat_region
/note="Charliel repeat: matches 1718..2101 of consensus"
1292..1654
repeat_region
/note="THE1B repeat: matches 1..364 of consensus"
1655..1995
repeat_region
/note="Charliel repeat: matches 2101..2527 of consensus"
2071..2202
repeat_region
/note="AluJdo/FRAM repeat: matches 147..281 of consensus"
2212..2350
repeat_region
/note="L1IMP2 repeat: matches 5995..6134 of consensus"
2428..2784
repeat_region
/note="L1MA5 repeat: matches 5925..6300 of consensus"
2788..2878
repeat_region
/note="Charliel repeat: matches 2513..2609 of consensus"
2879..3189
repeat_region
/note="AluY repeat: matches 1..311 of consensus"
3198..3493
repeat_region
/note="AluX repeat: matches 1..309 of consensus"
3494..3657
repeat_region
/note="Charliel repeat: matches 2616..2759 of consensus"
3743..4037
repeat_region
/note="AluX repeat: matches 1..296 of consensus"
5207..5261
/note="L2 repeat: matches 2683..2737 of consensus"

repeat_region
5349..5726
/note="MLR1A1 repeat: matches 1..365 of consensus"
7228..7418
repeat_region
/note="MLR1J repeat: matches 30..244 of consensus"
7546..7690
repeat_region
/note="MLR1J repeat: matches 342..488 of consensus"
7981..8260
repeat_region
/note="AluJo repeat: matches 1..298 of consensus"
10155..10361
repeat_region
/note="MER20 repeat: matches 6..218 of consensus"
10600..10978
repeat_region
/note="MLR1B repeat: matches 1..390 of consensus"
11362..11491
repeat_region
/note="L1ME3A repeat: matches 6025..6154 of consensus"
12192..12242
repeat_region
/note="LTR30 repeat: matches 1..52 of consensus"
12243..12783
repeat_region
/note="LTR12 repeat: matches 38..2212 of consensus"
13645..14352
repeat_region
/note="L1MC5 repeat: matches 7190..7909 of consensus"
14355..14918
repeat_region
/note="MLR1B repeat: matches 1..568 of consensus"
14922..15138
repeat_region
/note="L1MC5 repeat: matches 6966..7184 of consensus"
15394..15485
repeat_region
/note="L1M4 repeat: matches 5557..5647 of consensus"
15799..15853
repeat_region
/note="MSTA repeat: matches 373..426 of consensus"
15854..15893
repeat_region
/note="20 copies 2 mer tt 77% conserved"
15895..17669
repeat_region
/note="L1PA2 repeat: matches 4369..6146 of consensus"
17670..17970
repeat_region
/note="MSTA repeat: matches 9..375 of consensus"
18571..18624
repeat_region
/note="27 copies 2 mer ac 100% conserved"
18651..18823
repeat_region
/note="L2 repeat: matches 1273..1436 of consensus"
19824..19972
repeat_region
/note="MIR repeat: matches 112..261 of consensus"
20171..20306
repeat_region
/note="L1MA9 repeat: matches 6165..6302 of consensus"
20310..20623
repeat_region
/note="AluJo repeat: matches 1..306 of consensus"
20637..20865
repeat_region
/note="L1MA9 repeat: matches 5783..6030 of consensus"
21055..21274
repeat_region
/note="L2 repeat: matches 1332..1598 of consensus"
21309..21595
repeat_region
/note="ALUSC repeat: matches 1..287 of consensus"
22367..22675
repeat_region
/note="HAL1 repeat: matches 483..800 of consensus"
23702..23775
repeat_region
/note="L2 repeat: matches 2611..2699 of consensus"
27142..27297
repeat_region
/note="FLAM C repeat: matches 2..122 of consensus"
29141..29224
repeat_region
/note="MIR repeat: matches 56..153 of consensus"
30948..30993
repeat_region
/note="23 copies 2 mer tg 87% conserved"
32111..32154
repeat_region
/note="22 copies 2 mer ta 79% conserved"
32856..33026
repeat_region
/note="MIR repeat: matches 15..182 of consensus"
33169..33235
repeat_region
/note="MER5A repeat: matches 124..189 of consensus"
34625..34728
repeat_region
/note="MIR repeat: matches 78..184 of consensus"
35212..35409
repeat_region
/note="MER3 repeat: matches 3..208 of consensus"
35809..36116
repeat_region
/note="AluJb repeat: matches 1..302 of consensus"
36126..36197

```
repeat_region /note="MIR repeat: matches 73. .144 of consensus"
37119. .37244
/note="WERSA repeat: matches 73. .189 of consensus"
37713. .37765
/note="WSTD repeat: matches 1. .53 of consensus"
37766. .38064
/note="AluSq repeat: matches 1. .298 of consensus"
38065. .38409
/note="WSTD repeat: matches 53. .392 of consensus"
38512. .38630
/note="MIR repeat: matches 36. .155 of consensus"
38651. .38776
/note="L2 repeat: matches 1354. .1486 of consensus"
39209. .39304
/note="L2 repeat: matches 1818. .1924 of consensus"
40099. .40275
/note="WERSA repeat: matches 1. .168 of consensus"
41200. .41267
/note="34 copies 2 mer ac 80% conserved"
42055. .42551
/note="WL2CA repeat: matches 1. .503 of consensus"
43484. .43529
/note="23 copies 2 mer ca 97% conserved"
45011. .45184
/note="MIR repeat: matches 15. .188 of consensus"
45185. .45740
/note="WL2D repeat: matches 1. .551 of consensus"
45741. .45815
/note="MIR repeat: matches 188. .262 of consensus"
46651. .46956
/note="WER2 repeat: matches 1. .345 of consensus"
47341. .47633
/note="AluY repeat: matches 1. .293 of consensus"
47641. .48099
/note="WL2D repeat: matches 1. .505 of consensus"
48139. .48307
/note="AluCb repeat: matches 130. .301 of consensus"
48373. .48498
/note="AluCb repeat: matches 1. .127 of consensus"
49476. .49630
/note="MIR repeat: matches 22. .190 of consensus"
50545. .50877
/note="L2 repeat: matches 1782. .2161 of consensus"
50946. .51187
repeat_region

Query Match 28.6%; Score 286; DB 9; Length 129866;
Best Local Similarity 84.3%; Pred. No. 1.8e-62;
Matches 322; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 1 AACAGAGAACATCTCGCGCGGCGTGTGGCTCACACCTGTATCCCAACACTTT 60
Db 3209 AATAAGATAAACCATTTGCGCGCGGCGCGGTGCTCACCGCTGTATCCAGCACTTT 3150

Qy 61 GGGAGGCGGCGGCGGTGATCACTAGTGTAGGAGATCGAGACCACTCTGCTGACACGG 120
Db 3149 GGGAGGCGGAGCGGCGGATCACGAGTCAAGAGATCGAGACCACTCTGCTTAACCG 3090

Qy 121 TGAACCTCTCTTACTATAAAACACAAAAAATTAGCCGGCGGTGTGGCAGGCGCCTGT 180
Db 3089 TGAACCTCTCTTACTATAAAATACAAAAAATTAGCCGGCGGTGTGAGCGCGCCTGT 3030

Qy 181 AGTCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTCAACCCGGAGCGCGAATTT 240
Db 3029 AGTCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCGTGAACCCGGAGCGCGAGCTT 2970

Qy 241 GCAGTGAGCCGAGGTTCACCACTCACTCCAGCTGGGCAACACAGTGAGATCCGCTCT 300
Db 2969 GCAGTGAGCCGAGATCGCACCACTCACTCCAGCTGGGCAACCCGGAGCGCGAGCTT 2910

Qy 301 CAAAAAAGAAAGAAAGAAATACATCTCTTCAATGTTGGGATTTAAAC 360
Db 2909 CAAAAAAGAAAGAAAGAAAGAAATACATCTCTTCAATGTTGGGATTTAAAC 2850
```

```
Qy 361 TATTAGCCTTTCTGTAGCAA 382
Db 2849 TAAATACATATTTCTTATAAA 2828

RESULT 12
AL359177/c
LOCUS AL359177 230141 bp DNA linear HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-413B16, 16 unordered pieces.
ACCESSION AL359177
VERSION AL359177.3 GI:12331067
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Burton, J.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonsrequest@sanger.ac.uk
On Jan 22, 2001 this sequence version replaced gi:9797929.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA413B16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 127350 bases at least Q40
Consensus quality: 132189 bases at least Q30
Consensus quality: 134528 bases at least Q20
Insert size: 228641; sum-of-contigs
Insert size: 152695; 8.1% error; agarose-fp
Quality coverage: 2.04x in Q20 bases; sum-of-contigs Quality
coverage: 3.11x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 3155: contig of 3155 bp in length
* 3156
* 3255: gap of 100 bp
* 3256
* 10055: contig of 800 bp in length
* 10056
* 10155: gap of 100 bp
* 10156
* 16833: contig of 6678 bp in length
* 16834
* 16933: gap of 100 bp
* 16934
* 36991: contig of 20058 bp in length
* 36992
* 37091: gap of 100 bp
* 37092
* 46051: contig of 8960 bp in length
* 46052
* 46151: gap of 100 bp
* 46152
* 58091: contig of 11940 bp in length
* 58092
* 58191: gap of 100 bp
* 58192
* 64171: contig of 5980 bp in length
* 64172
* 64271: gap of 100 bp
* 64272
* 70251: contig of 5980 bp in length
* 70252
* 70351: gap of 100 bp
* 70352
* 73351: contig of 3000 bp in length
* 73352
* 73451: gap of 100 bp
* 73452
* 73531: contig of 5980 bp in length
* 73532
* 79432
* 79531: gap of 100 bp
* 79532
* 85511: contig of 5980 bp in length
* 85512
* 85611: gap of 100 bp
* 85612
* 100531: contig of 14920 bp in length
```

* 100532 100631: gap of 100 bp
* 100632 106611: contig of 5980 bp in length
* 106612 106711: gap of 100 bp
* 106712 125459: contig of 18748 bp in length
* 125460 125559: gap of 100 bp
* 125560 134519: contig of 8960 bp in length
* 134520 134619: gap of 100 bp
* 134620 230141: contig of 95522 bp in length.

FEATURES

Location/Qualifiers
1..230141

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-413B16"
/clone_lib="RPC1-11.2"

misc_feature

1..3155
/note="assembly_fragment:00110"

misc_feature

3256..10055
/note="assembly_fragment:00856"

misc_feature

10156..16833
/note="assembly_fragment:01004"

misc_feature

16934..36991
/note="assembly_fragment:01205"

misc_feature

37092..46051
/note="assembly_fragment:01284"

misc_feature

46152..58091
/note="assembly_fragment:01287"

misc_feature

58192..64171
/note="assembly_fragment:01291"

misc_feature

64272..70251
/note="assembly_fragment:01293"

misc_feature

70352..73351
/note="assembly_fragment:01295"

misc_feature

73452..79431
/note="assembly_fragment:01300"

misc_feature

79532..85511
/note="assembly_fragment:01302"

misc_feature

85612..100531
/note="assembly_fragment:01305"

misc_feature

100632..106611
/note="assembly_fragment:01306"

misc_feature

106712..125459
/note="assembly_fragment:01310"

misc_feature

125560..134519
/note="assembly_fragment:01312"

misc_feature

134620..230141
/note="assembly_fragment:01255"

clone_end:17
vector_side:right

ORIGIN

Query Match 28.5%; Score 285.4; DB 2; Length 230141;
Best Local Similarity 90.7%; Pred. No. 2.7e-62;
Matches 304; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 21 GSCCGGGCGGTGGTGTACACCTGTAATCCCAACACTTTGGAGGCGGAGCGGGTGA 80
Db 70245 GSCCGGGCGGTGGTGTACACCTGTAATCCCAACACTTTGGAGGCGGAGCGGGTGA 70186
QY 81 TCATAGTTCAGGAGATGAGACCAATCTTGGCTGACAGGTGAACCCCTGCTCTACTAA 140
Db 70185 TCACGAGTTCAGGAGATGAGACCAATCTTGGCTGACAGGTGAACCCCTGCTCTACTAA 70126
QY 141 AAACACAAAATTAGCGGGCGGTGGTGGCAGCGCCCTGTAGTCCACCTACTCTCGGAGG 200
Db 70125 AAATACAAAATTAGCGGGCGGTGGTGGCAGCGCCCTGTAGTCCACCTACTCTCGGAGG 70066
QY 201 CTGAGCAGGAGAAATGGCTGAACCCGGGAGCGGAACTTGAGTGAGCCGAGGTTGCAC 260
Db 70065 CTGAGCAGGAGAAATGGCTGAACCCGGGAGCGGAGCTTGAGTGAGCCGAGGTTGCAC 70006
QY 261 CACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGTCTCAAAAAAAGAAAAAG 320

Db 70005 CACTGCACTCCAGCTGGCGCACAGAGAGACCTCTCAAAAAAAGAAAAA 69946
QY 321 AAAGAAATCACATCTCATTCAGTGGTGGCATTT 355
Db 69945 AAAAAAATTTCTCCCACTTATAGTGGTGCCTAT 69911
RESULT 13
AC064871
LOCUS AC064871 194197 bp DNA linear PRI 05-DEC-2002
DEFINITION Homo sapiens BAC clone RP11-598C21 from 2, complete sequence.
ACCESSION AC064871
VERSION AC064871.6 GI:15431251
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 194197)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 194197)
AUTHORS Belter, E., Haakensen, W., Hawkins, M., Elliott, G., Dixon, R. and Doebber, A.
TITLE The sequence of Homo sapiens BAC clone RP11-598C21
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 194197)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 194197)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 194197)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 194197)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 5, 2001 this sequence version replaced gi:15022768.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics

Center project name: H_NH0598C21

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Oseogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenoh,M., Catanese,J.U. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-60F10, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-598C21; actual end is at base position 194197 of RP11-598C21.

Data from AC079249 was used to finish this clone, AC064871. In AC06487, the region from 90272 to 90395 is only represented by data from AC079249. Polymorphisms exist between RP11-598C21 and RP11-60F10.

FEATURES

source

Location/Qualifiers

1..194197
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"

/clone="RP11-598C21"

/clone_lib="RPCI-11"

70..328

/rpt_family="MER1_type"

329..471

/rpt_family="MIR"

490..773

/rpt_family="Alu"

893..1028

/note="match to EST BF991908 (NID:g12398231)"

893..1028

/note="match to EST BG120186 (NID:g12613708)"

893..1028

/note="match to EST BG426467 (NID:g13332973)"

893..1028

/note="match to EST BI093191 (NID:g14511521)"

893..1028

/note="match to EST N31469 (NID:g1151868) yx55g03.r1"

893..1028

/note="similar to Homo sapiens EST BF242272 (NID:g1156199)"

893..1019

/note="match to EST N44074 (NID:g1182602) yy30h01.r1"

916..1028

/note="similar to Mus musculus EST BF140223 (NID:g10979263)"

986..1028

/note="match to EST AU133346 (NID:g10993885)"

1082..1103

/rpt_family="AT-rich"

1242..1568

/note="match to EST AL119168 (NID:gs925067)"

1476

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

1892..2031

/note="match to EST AU133346 (NID:g10993885)"

1892..2031

/note="match to EST BG120186 (NID:g12613708)"

1892..2031

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

repeat_region

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

repeat_region

repeat_region

repeat_region

misc_feature

misc_feature

misc_feature

misc_feature

/note="match to EST BG426467 (NID:g13332973)"

1892..2031

/note="match to EST BI093191 (NID:g14511521)"

1892..2031

/note="match to EST N31469 (NID:g1151868) yx55g03.r1"

1892..2031

/note="similar to Homo sapiens EST BF242272 (NID:g1156199)"

1892..2031

/note="similar to Mus musculus EST BF140223 (NID:g10979263)"

1892..1957

/note="match to EST BF991908 (NID:g12398231)"

1900..2031

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

1971..2031

/note="match to EST AU125389 (NID:g10950105)"

2576..2599

/rpt_family="(TTTTC)n"

2593..3011

/note="match to EST AA179659 (NID:g1761873) zp54f10.s1"

3023..3128

/rpt_family="L2"

3287..3539

/rpt_family="MER1_type"

3558..3597

/rpt_family="L2"

3757..3815

/rpt_family="L1"

6162..6191

/rpt_family="(TAAA)n"

6594..6713

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

6594..6713

/note="match to EST AU125389 (NID:g10950105)"

6594..6713

/note="match to EST AU133346 (NID:g10993885)"

6594..6713

/note="match to EST BG120186 (NID:g12613708)"

6594..6713

/note="match to EST BG426467 (NID:g13332973)"

6594..6713

/note="match to EST BI093191 (NID:g14511521)"

6594..6713

/note="similar to Homo sapiens EST BF242272 (NID:g1156199)"

6594..6713

/note="similar to Mus musculus EST BF140223 (NID:g10979263)"

6594..6614

/note="match to EST N31469 (NID:g1151868) yx55g03.r1"

7931..8013

/rpt_family="L2"

8280..8370

/rpt_family="ERVL"

8471..8516

/rpt_family="AT-rich"

9107..9243

/note="match to EST AW375848 (NID:g6880502)"

9111..9243

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

9111..9243

/note="match to EST AU125389 (NID:g10950105)"

9111..9243

Query Match 28.5%; Score 285; DB 9; Length 194197;

Best Local Similarity 87.4%; Pred.No.3.4e-62;

Matches 312; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 2 ACAGAGAAAGACATCTCGCCGGCGGTGGTGGCTACACCTTAATCCCAACATTG 61

Db 15529 AAGAGATAAAAGATATCGCGCGCGGTGGTGGCTACACCTTAATCCCAACATTG 15588

repeat_region 33209. .33586
/note="WLR1E repeat: matches 182. .567 of consensus"
repeat_region 33588. .33752
/note="WLR1F repeat: matches 261. .438 of consensus"
repeat_region 33777. .34258
/note="MER74B repeat: matches 81. .552 of consensus"
repeat_region 34300. .34522
/note="AluY repeat: matches 82. .298 of consensus"
repeat_region 34545. .34656
/note="WLR1F repeat: matches 431. .541 of consensus"
repeat_region 35792. .36598
/note="L1M4 repeat: matches 4252. .5114 of consensus"
repeat_region 37988. .38358
/note="THR1C repeat: matches 1. .371 of consensus"
repeat_region 41148. .41248
/note="WIR repeat: matches 51. .164 of consensus"
repeat_region 41315. .41974
/note="L2 repeat: matches 2099. .2748 of consensus"
repeat_region 43058. .43631
/note="MER44C repeat: matches 72. .728 of consensus"
repeat_region 43758. .43862
/note="L2 repeat: matches 2574. .2688 of consensus"
repeat_region 43952. .44257
/note="AluSC repeat: matches 1. .306 of consensus"
repeat_region 44863. .44927
/note="L2 repeat: matches 2605. .2668 of consensus"
repeat_region 44920. .44992
/note="L2 repeat: matches 2638. .2710 of consensus"
repeat_region 45521. .45625
/note="L2 repeat: matches 2606. .2710 of consensus"
repeat_region 46736. .47270
/note="WLR1F repeat: matches 17. .541 of consensus"
repeat_region 47361. .47612
/note="WLR2CB repeat: matches 1. .271 of consensus"
repeat_region 49013. .49271
/note="AluX repeat: matches 12. .276 of consensus"
repeat_region 50091. .50819
/note="LTR1 repeat: matches 33. .779 of consensus"
repeat_region 51674. .51843
/note="L2 repeat: matches 2539. .2710 of consensus"
repeat_region 54558. .54692
/note="L2 repeat: matches 2559. .2699 of consensus"
repeat_region 54687. .54887
/note="L2 repeat: matches 2205. .2417 of consensus"
repeat_region 55357. .55656
/note="AluSC repeat: matches 1. .298 of consensus"
repeat_region 59818. .60015
/note="L1 repeat: matches 4599. .4800 of consensus"
repeat_region 60016. .60327
/note="AluY repeat: matches 1. .311 of consensus"
repeat_region 60328. .60827
/note="L1 repeat: matches 4118. .4599 of consensus"
repeat_region 62797. .62917
/note="L1P47 repeat: matches 6023. .6143 of consensus"
repeat_region 63385. .63724
/note="WLR1B repeat: matches 1. .333 of consensus"
repeat_region 63726. .63844
/note="7 copies 17 mer 66% conserved"
repeat_region 63732. .64064
/note="9 copies 37 mer 65% conserved"
repeat_region 63810. .64169
/note="90 copies 4 mer tata 54% conserved"
repeat_region 63811. .64142
/note="166 copies 2 mer at 54% conserved"
repeat_region 64329. .64413
/note="5 copies 17 mer 71% conserved"
repeat_region 66588. .72711
/note="L1P44 repeat: matches 3. .6144 of consensus"
repeat_region 73548. .73599
/note="13 copies 4 mer gata 80% conserved"
repeat_region 73828. .74025
/note="MER2 repeat: matches 1. .200 of consensus"
repeat_region 74220. .75339

/note="L2 repeat: matches 1625. .2746 of consensus"
repeat_region 77029. .77064
/note="MIR repeat: matches 156. .191 of consensus"
repeat_region 77128. .77406
Query Match 28.4%; Score 284.2; DB 9; Length 98868;
Best Local Similarity 79.5%; Pred. No. 5.1e-62;
Matches 349; Conservative 0; Mismatches 88; Indels 2; Gaps 1;
Qy 7 AGAAGCACATCTCGGCGGCGTGGCTGCACCTGTAAATCCCAACACTTTGGGAGG 66
Db 8064 AAAAAATTTTCTCGGCGGCGGCTGCACCTGTAAATCCCAACACTTTGGGAGG 8005
Qy 67 CCGAGGCGGTGGATCACTAGGTGAGGATGAGACCATCTCTGGCTGCACCGTGAAC 126
Db 8004 CCGAGGCGGCGGATCAGGAGTCAAGGATCGAGACCATCTCTGGCTGCACCGTGAAC 7945
Qy 127 CTGTCTCTACTAAAAACAAAAATTAGCGGCGTGTGGCAGGCGCTGTAGTCCC 186
Db 7944 CCGTCTCTATAAAAAATAGCCGGCGTGTGGCGGCGCTGTAGTCCC 7885
Qy 187 AGCTACTCGGAGGCTGAGCAGGAGATGCCCTGAACCGGAGCGGAACTTGCAGTG 246
Db 7884 AGCTACTCGGAGGCTGAGCAGGAGATGCCCTGAACCGGAGCGGAGCTTGCAGTG 7825
Qy 247 AGCGAGGTTGCACCACTGCACCTCCAGCTGGGCAACACAGTGAAGTCCGTCTCAAAAA 306
Db 7824 AGCGAGATTGCGCCACTGCATCCAGCTGGGCGAGAGAGATCCGTCTCAAAAA 7765
Qy 307 AAAAAAGAAAAAGAAATCATCTCATTCAGTGTGGCATTTAAACTATTTA 366
Db 7764 AAAAAAGAAAAAGAAATATAT- -TTTCTCAGGTGTAGTTAAGGTAGCACA 7707
Qy 367 GCCTTCTGTAGCAAGTTAGTATCTGTCTTTCAGACCTCAAGGTGTTTTTCTTT 426
Db 7706 AGCTCTAAATATTCAAGTTATTATCTTGTATTCAAGAAATATAATATTCAAGTTATT 7647
Qy 427 GTTTTTCATACCGGTGTG 445
Db 7646 ATATTCTTAACTTTGAG 7628
RESULT 15
AC012462/c
LOCUS AC012462 143947 bp DNA linear PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-566P21 from 2, complete sequence.
AC012462
ACCESSION AC012462.5 GI:14589768
VERSION AC012462.5
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 143947)
AUTHORS Harris,A., Cotton,M. and Doebber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-566P21
JOURNAL Unpublished (2001)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 143947)
AUTHORS Harris,A., Cotton,M. and Doebber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-566P21
JOURNAL Unpublished (2001)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 3 (bases 1 to 143947)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 143947)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 143947)
Waterston,R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 143947)
Waterston,R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 3, 2001 this sequence version replaced gi:13518276.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0566P21

misc_feature	1. .113	/note="similar to EST AU138563 (NID:g11000084)"
misc_feature	1. .113	/note="match to EST AV752678 (NID:g10910526)"
misc_feature	1. .113	/note="match to EST AW068021 (NID:g5023019) cn20a08.x1"
misc_feature	1. .113	/note="match to EST AW608520 (NID:g7313261)"
misc_feature	1. .113	/note="match to EST BF751233 (NID:g12077909)"
misc_feature	1. .113	/note="match to EST BG163577 (NID:g12670280)"
misc_feature	1. .113	/note="match to EST BG260847 (NID:g12770663)"
misc_feature	1. .113	/note="similar to Bos taurus EST AV614913 (NID:g9750583)"
misc_feature	1. .113	/note="similar to EST BG897400 (NID:g14307641)"
misc_feature	1. .113	/note="similar to EST BG924585 (NID:g14319108)"
misc_feature	1. .113	/note="similar to EST BG927121 (NID:g14321644)"
misc_feature	1. .113	/note="similar to Mus musculus EST AL364292 (NID:g9691680)"
misc_feature	8. .113	/note="match to EST AI307037 (NID:g6497564)"
misc_feature	45. .113	/note="match to EST AU119133 (NID:g10934368)"
misc_feature	57. .113	/note="match to EST AW847780 (NID:g7943297)"
misc_feature	59. .113	/note="match to EST AU137394 (NID:g10997933)"
misc_feature	63. .113	/note="match to EST AU118498 (NID:g10933555)"
misc_feature	78. .113	/note="match to EST BE934096 (NID:g10460172)"
misc_feature	78. .113	/note="similar to EST BG993547 (NID:g14397617)"
misc_feature	84. .113	/note="match to EST AW608477 (NID:g7313218)"
misc_feature	93. .113	/note="match to EST BF339928 (NID:g11286390)"
misc_feature	173. .177	/note="similar to Homo sapiens EST AW058318 (NID:g59333957) wx18h04.x1"
misc_feature	255. .282	/note="match to EST AI378776 (NID:g4188629) tc18h12.x1"
repeat_region	261. .304	/rpt_family=" (TTTA)n"
repeat_region	284. .599	/rpt_family="Alu"
misc_feature	290. .307	/note="match to EST AI271477 (NID:g3890644) qi19f12.x1"
repeat_region	450. .481	/rpt_family=" (T)n"
misc_feature	465. .941	/note="match to EST AI333596 (NID:g4070155) qq06h09.x1"
repeat_region	622. .644	/rpt_family="AT_rich"
misc_feature	925. .1258	/note="match to EST H77493 (NID:g1055582) yu09b11.r1"
misc_feature	969. .1339	/note="match to EST AA953560 (NID:g3117707) om95g02.s1"
misc_feature	1019. .1462	/note="match to EST D63177 (NID:g9966846)"
misc_feature	1319. .1500	/note="match to EST BF326027 (NID:g11296775)"
misc_feature	1722	/note="match to EST AA852317 (NID:g2940910)"
misc_feature	1864. .1899	/note="match to EST AV652269 (NID:g9873283)"
misc_feature	1869. .1899	

```
misc_feature /note="match to EST AV652315 (NID:g9873329)"
1926..2448
misc_feature /note="match to EST BE169125 (NID:g8631846)"
1951..2289
misc_feature /note="match to EST AV652269 (NID:g9873283)"
1951..2289
misc_feature /note="match to EST AV652315 (NID:g9873329)"
2036..2037
misc_feature /note="match to EST BE812565 (NID:g10244903)"
2140
misc_feature /note="match to EST AV653321 (NID:g9874335)"
2219..2347
misc_feature /note="match to EST BF749100 (NID:g12075776)"
2322..2323
misc_feature /note="match to EST AW371903 (NID:g6876557)"
2345..2510
misc_feature /note="match to EST BG115942 (NID:g12609448)"
2346..2515
misc_feature /note="match to EST BF751233 (NID:g12077909)"
2346..2511
misc_feature /note="match to EST AW68021 (NID:g6023019) cn20a08.xl"
2346..2510
misc_feature /note="match to EST AI907037 (NID:g6497564)"
2346..2510
misc_feature /note="match to EST AU117358 (NID:g10932320)"

Query Match 28.3%; Score 283.4; DB 9; Length 143947;
Best Local Similarity 90.7%; Pred. No. 8.4e-62;
Matches 302; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 60
Db 60950 AAAAGTTGAAGAAGAAAGATCGCCGCGCGCGTGGCTCACCCCTGTATATCCAGCACTTT 60891

Qy 61 GGGAGCGCGAGCGCGGTGGATCACTAGGTCAAGAGATGGAGACCATCTGGCTGACACGG 120
Db 60890 GGGAGGCCAAGCGCGGTGGATCAAGGTCAAGAGATCGAGACCATCTGGCTAACACGG 60831

Qy 121 TGAACCCCTGTCTTACTAAAAACACAAAAATTAGCGCGCGGTGGTGGAGCGCCTGT 180
Db 60830 TGAACCCCGTCTTACTAAAAATACAAAAATTAGCGCGCGGTGGTGGCGGCGCCTGT 60771

Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCGGGAGCGGAACTT 240
Db 60770 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCGGGAGCGAGACTT 60711

Qy 241 GCAGTGAGCCGAGTTGCACCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCT 300
Db 60710 GCAGTGAGCCGAGATTGGCCACTGCACTCCAGCTGGGTGACAGAGTGAGACTCCGCTCT 60651

Qy 301 CAAAAAAGAAAAAGAAAAAGAAATACCA 333
Db 60650 CAAAAAAGAAAAAAGAAAAAAGAAAAAGA 60618
```

Search completed: July 4, 2005, 09:55:37
Job time : 2892.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 352.271 Seconds

(without alignments)
16821.327 Million cell updates/sec

Title: US-09-936-271c-13_COPY_8000_9000

Perfect score: 1001

Sequence: 1 aacagaagaagcaccatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	3	AA95905
2	1001	100.0	11570	12	ADK52482
3	1001	100.0	11570	13	ADR72623
4	1001	100.0	11570	13	ADR72875
5	283.4	28.3	31749	4	AAK72959
6	280	28.0	78925	3	AAK89888
7	279.4	27.9	96599	9	AAK57703
8	279.4	27.9	96599	9	ADA02672
9	279.4	27.9	96599	10	ADB72410
10	278.4	27.9	96599	10	ADE95920
11	278.8	27.9	77781	10	ADL15049
12	278.8	27.9	110300	13	ADS36499
13	278.2	27.8	3407	4	AAK90223
14	278.2	27.8	3407	5	AAK39852
15	278.2	27.8	3407	9	ADB32812
16	277.6	27.7	347	4	AAK81654
17	277.6	27.7	5115	4	AAK83483
18	277.6	27.7	49777	4	AAK75029
19	277.6	27.7	49777	4	AAK76214
20	277.2	27.7	31952	4	AAK89370

c	22	276.8	27.7	114793	4	AA080215	AA080215 Human gen
c	22	276.6	27.6	119501	12	ADI29095	ADI29095 Human MAR
c	23	276.4	27.6	31474	4	AA05461	AA05461 Human rep
c	24	276.4	27.6	31474	4	ABL98314	ABL98314 Human tes
c	25	276.4	27.6	204621	11	ACN44486	ACN44486 Human gen
c	26	276.2	27.5	100445	13	ABD33179	ABD33179 Human can
c	27	275.2	27.5	41104	6	AA036260	AA036260 Human G-p
c	28	274.8	27.5	2328	4	AA03890	AA03890 Human rep
c	29	274.6	27.4	1400	12	ADH70101	ADH70101 Human Vbe
c	30	274.6	27.4	2325	4	AA03891	AA03891 Human rep
c	31	274.6	27.4	3608	4	AAK83192	AAK83192 Human imm
c	32	274.6	27.4	3608	4	AAK74891	AAK74891 Human imm
c	33	274.6	27.4	3608	4	AAK67271	AAK67271 Human imm
c	34	274.6	27.4	110000	12	ADH69807_1	Continuation (2 of
c	35	274.6	27.4	177380	8	ACF62751	ACF62751 Cancer ba
c	36	274.6	27.4	177380	8	ADB20870	ADB20870 MRPI base
c	37	274.6	27.4	177380	10	ADB87959	ADB87959 Human UGT
c	38	274.6	27.4	177380	10	ADB96942	ADB96942 Human MDR
c	39	274.6	27.4	177380	10	ADB92133	ADB92133 Human MDR
c	40	274.6	27.4	267156	6	ABL68560	ABL68560 Kidney ca
c	41	274.4	27.4	465	5	ABV35226	ABV35226 Human pro
c	42	274.2	27.4	14543	6	ABK15798	ABK15798 Human von
c	43	274	27.4	1992	4	AAK85653	AAK85653 Human imm
c	44	274	27.4	1995	4	AAK85652	AAK85652 Human imm
c	45	274	27.4	17181	4	AAK72508	AAK72508 Human imm

ALIGNMENTS

RESULT 1

AAA95905
ID AAA95905 standard; DNA; 11570 BP.
XX
AC AAA95905;
DT 02-FEB-2001 (first entry)
XX
DE Human KLK-L2 gene.
XX
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KW kallikrein-like protein; serine protease; cytostatic; cancer;
KW prostrate cancer; ds.
XX
OS Homo sapiens.
XX
PN WO200053776-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-CA000258.
XX
PR 11-MAR-1999; 99US-0124260P.
PR 01-APR-1999; 99US-0127386P.
PR 21-JUL-1999; 99US-0144919P.
PA (MOUN) MOUNT SINAI HOSPITAL.
XX
PI Aac89888 Human FN
PI Aal57703 Human pro
XX
DR Ado2672 Human PRL
DR Adb72410 Human PRL
DR Ade95920 Human PRL
XX
PT Adl15049 Human mel
PT AdS36499 Human aut
XX
PS Aak90223 Human dig
XX Aas39852 Genomic s
XX Adb32812 Human nov
CC Aak81654 Human imm
CC Aak83483 Human imm
CC Aak75029 Human imm
CC Aak76214 Human imm
CC Aak89370 Human dig

CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

XX
SQ Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 3; Length 11570;
Best Local Similarity 100.0%; Pred. No. 2.9e-239;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTAAATCCCAACACTTT 60
Db 8000 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTAAATCCCAACACTTT 8059

Qy 61 GGGAGCGCGAGCGGGTGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 120
Db 8060 GGGAGCGCGAGCGGGTGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 8119

Qy 121 TGAACCTGTCTCTACTAAAAACACAAAAAATTAGCGCGCGTGGTGGAGGCGCTGT 180
Db 8120 TGAACCTGTCTCTACTAAAAACACAAAAAATTAGCGCGCGTGGTGGAGGCGCTGT 8179

Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCGGAGGCGGAACCTT 240
Db 8180 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCGGAGGCGGAACCTT 8239

Qy 241 GCAGTGAGCGAGGTGACCACTGCATCTCAGCCTGGGCAACACAGTGAGACTCCGTCT 300
Db 8240 GCAGTGAGCGAGGTGACCACTGCATCTCAGCCTGGGCAACACAGTGAGACTCCGTCT 8299

Qy 301 CAAAAAAGAAAAAGAAAAAGAAATCAATCATCTCATTTCAAGTGGTGGCAATTAAC 360
Db 8300 CAAAAAAGAAAAAGAAAAAGAAATCAATCATCTCATTTCAAGTGGTGGCAATTAAC 8359

Qy 361 TATTAGCTTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGTGGTTTTT 420
Db 8360 TATTAGCTTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGTGGTTTTT 8419

Qy 421 TTGTTGTTTTTTCATACCGGTGTGTGTCTGGGTGGCCACTAAAAGCTACAGCAAG 480
Db 8420 TTGTTGTTTTTTCATACCGGTGTGTGTCTGGGTGGCCACTAAAAGCTACAGCAAG 8479

Qy 481 AAATAAACAACCTACACAATATCTATACCAATAGTATATAAATAATAGCATCTGGCTA 540
Db 8480 AAATAAACAACCTACACAATATCTATACCAATAGTATATAAATAATAGCATCTGGCTA 8539

Qy 541 ATTGCTGACACTGTTTTAAGTGGTTTGCATGCTCAGCTCATTAACCTCATTTACCTGTT 600
Db 8540 ATTGCTGACACTGTTTTAAGTGGTTTGCATGCTCAGCTCATTAACCTCATTTACCTGTT 8599

Qy 601 ATTATTGSCCTATTTTTAAACAAAGGAGGCAAGGCTCAGAGCTGTTAACTAACAGCTC 660
Db 8600 ATTATTGSCCTATTTTTAAACAAAGGAGGCAAGGCTCAGAGCTGTTAACTAACAGCTC 8659

Qy 661 TCAAAAGAAACCTGCGAGAGATATTAATTTAAAAAATAATAGAGAAATTAACCAACAA 720
Db 8660 TCAAAAGAAACCTGCGAGAGATATTAATTTAAAAAATAATAGAGAAATTAACCAACAA 8719

Qy 721 GAAAGTTGAAATTTAGAGGTACAGCAGTAAAGCTTGTGCTTTTGAACAGTCTGCT 780
Db 8720 GAAAGTTGAAATTTAGAGGTACAGCAGTAAAGCTTGTGCTTTTGAACAGTCTGCT 8779

Qy 781 ACTGGAAAAAGGCAAGCTTTGGCTTTTCTTAATTAATGATACCAAGCACTCTGTAATTCAT 840
Db 8780 ACTGGAAAAAGGCAAGCTTTGGCTTTTCTTAATTAATGATACCAAGCACTCTGTAATTCAT 8839

Qy 841 ATTTTGCATGATGTAAGTAAAGATGAAGCGGGTGAATGGACATGCGCAGTAATCCC 900
Db 8840 ATTTTGCATGATGTAAGTAAAGATGAAGCGGGTGAATGGACATGCGCAGTAATCCC 8899

Qy 901 AGCACTCTGGAGACTGAAGTGGAGATCACTTGGCTCAGGAGTTCAAGACCAGCTG 960
Db 8900 AGCACTCTGGAGACTGAAGTGGAGATCACTTGGCTCAGGAGTTCAAGACCAGCTG 8959

Qy 961 GGCACTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1001
Db 8960 GGCACTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9000

RESULT 2
ADK52482
ID ADK52482 standard; DNA; 11570 BP.
XX
AC ADK52482;
XX
DT 03-JUN-2004 (first entry)
XX Human kallikrein 5 encoding sequence.
XX kallikrein 5; cancer; Cytostatic; Immunostimulant; ds.
XX Homo sapiens.
XX WO2004021008-A2.
XX 11-MAR-2004.
XX 28-AUG-2003; 2003WO-CA001310.
XX 28-AUG-2002; 2002US-0407333P.
XX (MOUN) MOUNT SINAI HOSPITAL.
XX (YOUS/) YOUSEF G.
XX Diamandis EP;
XX WPI; 2004-239232/22.
XX P-PSDB; ADK52481.
XX Detecting kallikrein 5 associated with cancer, useful in diagnosing, monitoring, detecting, imaging and treating breast or ovarian carcinoma, comprises comparing the detected amount of kallikrein 5 in a sample with a standard sample.
XX Disclosure; SEQ ID NO 2; 59pp; English.
XX The present invention relates to detecting kallikrein 5 associated with breast or ovarian cancer in a patient comprises detecting in the sample kallikrein 5 and comparing the detected amount with an amount detected for a standard. The method is useful in detecting kallikrein 5 associated with breast or ovarian cancer in a patient. The methods and kits are useful in diagnosing, monitoring, detecting, imaging and treating breast or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for preventing and treating breast and ovarian cancer and for stimulating or enhancing antibody production or for inducing an immune response. The present sequence represents human kallikrein 5 encoding sequence.

XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 12; Length 11570;
Best Local Similarity 100.0%; Pred. No. 2.9e-239;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTAAATCCCAACACTTT 60
Db 8000 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTAAATCCCAACACTTT 8059

Qy 61 GGGAGCGCGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 120
Db 8060 GGGAGCGCGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 8119

Qy 121 TGAACCTGTCTCTACTAAAAACACAAAAAATTAGCGCGCGTGGTGGAGGCGCTGT 180

Db 8120 TGAACCCCTGCTCTACTAAAAACACAAAAAATTAGCCGGCGTGTGGCAGGCGCCTGT 8179
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGCGCTGAAACCCGGAGGCGGAATTT 240
Db 8180 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGCGCTGAAACCCGGAGGCGGAATTT 8239
Qy 241 GCAGTGAGCCGAGGTTGCACCACTGCACCTCAGGCTGGGCAACACAGTGAGACTCCGCTCT 300
Db 8240 GCAGTGAGCCGAGGTTGCACCACTGCACCTCAGGCTGGGCAACACAGTGAGACTCCGCTCT 8299
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATCAATCATCTCAATCAAGTGTGGCATTTAAAAAC 360
Db 8300 CAAAAAAGAAAAAGAAAAAGAAATCAATCATCTCAATCAAGTGTGGCATTTAAAAAC 8359
Qy 361 TATTTAGCCTTCTCTAGGCAAGGTTAGTATCTTGTGTTTTCAGACCTCAAGTGTGTTTT 420
Db 8360 TATTTAGCCTTCTCTAGGCAAGGTTAGTATCTTGTGTTTTCAGACCTCAAGTGTGTTTT 8419
Qy 421 TTGTTTGTGTTTTTCATACCGGTGTGTGTCTGGGTGTGGCCACTAAAAAGCTTACAAGCAAG 480
Db 8420 TTGTTTGTGTTTTTCATACCGGTGTGTGTCTGGGTGTGGCCACTAAAAAGCTTACAAGCAAG 8479
Qy 481 AAATAATAACAACATCAACAATCTAATACCAATAGTATATAAATAATATAGCATCTGGCTA 540
Db 8480 AAATAATAACAACATCAACAATCTAATACCAATAGTATATAAATAATATAGCATCTGGCTA 8539
Qy 541 ATTGCTGACACTGTGTTTAAAGTGTGTCAGCTCAGCTCAATTAATCAATTAACCTGTT 600
Db 8540 ATTGCTGACACTGTGTTTAAAGTGTGTCAGCTCAGCTCAATTAATCAATTAACCTGTT 8599
Qy 601 ATTATGCGCCCTATTATCAACAGGAGGCAAGGCTCAGAGCTTAACCTAAACAGGCTC 660
Db 8600 ATTATGCGCCCTATTATCAACAGGAGGCAAGGCTCAGAGCTTAACCTAAACAGGCTC 8659
Qy 661 TCAAAAAGAACTCTGACAGATATTAAATTTAAAAAATAATGAGAGAAATTAACCCACAA 720
Db 8660 TCAAAAAGAACTCTGACAGATATTAAATTTAAAAAATAATGAGAGAAATTAACCCACAA 8719
Qy 721 GAAAGTTGAAATTTAGAGGTACAGGCTAGCTTGTGCTTTGAAACAGTGTCTGCT 780
Db 8720 GAAAGTTGAAATTTAGAGGTACAGGCTAGCTTGTGCTTTGAAACAGTGTCTGCT 8779
Qy 781 ACTGGHAAAAGGCAAGTCTTGGCTTCTTAATTAATGATACCAAGCTCTGTAATTCAT 840
Db 8780 ACTGGHAAAAGGCAAGTCTTGGCTTCTTAATTAATGATACCAAGCTCTGTAATTCAT 8839
Qy 841 ATTTTGCATGATGTAAGTAAAGAAATGAAGCCGGTGCAATGGCAGTAAATCCC 900
Db 8840 ATTTTGCATGATGTAAGTAAAGAAATGAAGCCGGTGCAATGGCAGTAAATCCC 8899
Qy 901 AGCACTCTGGAGACTGAAGTGGGAGATCACTTGAAGTCAAGGTTCAAGCAGGCTG 960
Db 8900 AGCACTCTGGAGACTGAAGTGGGAGATCACTTGAAGTCAAGGTTCAAGCAGGCTG 8959
Qy 961 GGCAACTAAAAATTAAAAAATAAATACTAATGTTTTT 1001
Db 8960 GGCAACTAAAAATTAAAAAATAAATACTAATGTTTTT 9000

RESULT 3

ADNR72623

ID ADNR72623 standard; DNA; 11570 BP.

XX AC

ADNR72623;

XX DT 02-DEC-2004 (first entry)

XX DE Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.

XX KW kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;
XX KW cancer metastasis; chemotherapy; human; serine protease;
XX KW chromosome 19q13.4; KLK5; ds; gene.

XX

Homo sapiens.

XX Key Location/Qualifiers
PH CDS 2221..11247

FT /*tag= b

FT /product= "Human renal cell carcinoma-related kallikrein

FT 5 (hKS) protein"

FT 2221..2293

FT /*tag= a

FT /number= 1

FT 2294..4761

FT /*tag= c

FT /number= 1

FT 4762..5023

FT /*tag= d

FT /number= 2

FT 5024..5762

FT /*tag= e

FT /number= 2

FT 5763..6019

FT /*tag= f

FT /number= 3

FT 6020..6104

FT /*tag= g

FT /number= 3

FT 6105..6238

FT /*tag= h

FT /number= 4

FT 6239..11091

FT /*tag= i

FT /number= 4

FT 11092..11247

FT /*tag= j

FT /number= 5

WO2004077060-A2.

10-SEP-2004.

26-FEB-2004; 2004WO-CA000280.

27-FEB-2003; 2003US-0451382P.

(MOUN) MOUNT SINAI HOSPITAL.

Diamandis EP, Petraki CD;

WPI; 2004-662077/64.

P-PSDB; ADR72621.

Detecting kallikrein polypeptides/polynucleotides associated with renal cell carcinoma in a patient, for diagnosing/treating the disease, comprises detecting /identifying kallikrein polypeptides/polynucleotides in a sample.

Example 1; SEQ ID NO 3; 53pp; English.

The invention relates to a novel method for detecting kallikrein polypeptides, or the polynucleotides encoding them, associated with renal cell carcinoma. The method comprises obtaining a sample from a patient and detecting kallikrein polypeptides, or their encoding polynucleotides, where the kallikrein polypeptides are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The detected amounts of the kallikrein polypeptides are compared to standard amounts. The molecules of the invention demonstrate cytostatic activity whilst the methods and kit may be useful for detecting, characterising, preventing and treating renal cell carcinoma. Furthermore, the methods may be useful for evaluating the probability of the presence of malignant or pre-malignant cells and for detecting and quantitating tumour growth and cancer metastasis. Finally, the methods may be utilised to confirm the absence or removal of all tumour tissue following surgery, cancer chemotherapy and/or radiation therapy and to monitor cancer chemotherapy and tumour reappearance. The current sequence is that of the human

CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the
CC invention which encodes a secreted serine protease and is located at
CC chromosome 19q13.4.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 1001; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 2.9e-239;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACAGAGAAAGACATCTCGGCGGGGCTGGTGGCTCACACCTGTATCCCAACACTTT 60
Db AACAGAGAAAGACATCTCGGCGGGGCTGGTGGCTCACACCTGTATCCCAACACTTT 8059
Qy 61 GGGAGGCGGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGCTGCACACGG 120
Db GGGAGGCGGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGCTGCACACGG 8119
Qy 121 TGAACCCCTGTCTCTACTATAAAACACAAAAATTAGCCGGCGTGGTGCGAGGCGCCTGT 180
Db TGAACCCCTGTCTCTACTATAAAACACAAAAATTAGCCGGCGTGGTGCGAGGCGCCTGT 8179
Qy 181 AGTCCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTGAACCCGGGAGCGGAACTT 240
Db AGTCCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTGAACCCGGGAGCGGAACTT 8239
Qy 241 GCAGTGAGCGGAGGTGCACCACTGCATCTCAGCTCGGCAACACAGTGAGCTCCGCT 300
Db GCAGTGAGCGGAGGTGCACCACTGCATCTCAGCTCGGCAACACAGTGAGCTCCGCT 8299
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATCACTCATCTCAAGTGGTGGCAATTTAAAAAC 360
Db CAAAAAAGAAAAAGAAAAAGAAAAATCACTCATCTCAAGTGGTGGCAATTTAAAAAC 8359
Qy 361 TATTAGCTTCTGTAGGCAAGTTAGTAGTCTTTTTCAGACCTCAAGGTGTTTTT 420
Db TATTAGCTTCTGTAGGCAAGTTAGTAGTCTTTTTCAGACCTCAAGGTGTTTTT 8419
Qy 421 TTGTTTGTGTTTTTCATACCGGTGTGGTCTGGGTGGCCACTAAAAGCTTACAGCAAG 480
Db TTGTTTGTGTTTTTCATACCGGTGTGGTCTGGGTGGCCACTAAAAGCTTACAGCAAG 8479
Qy 481 AAATAAATAACAATACTAATAACCAATAGTATATAAATAATAGCATCTGGCTA 540
Db AAATAAATAACAATACTAATAACCAATAGTATATAAATAATAGCATCTGGCTA 8539
Qy 541 ATTGCTGACACTGTGTTTAAGTGGTTGATGCTCAGCTCATTAATCAATTTACCTGTT 600
Db ATTGCTGACACTGTGTTTAAGTGGTTGATGCTCAGCTCATTAATCAATTTACCTGTT 8599
Qy 601 ATTATTGGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCAGTTAACTAACAGGCTC 660
Db ATTATTGGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCAGTTAACTAACAGGCTC 8659
Qy 661 TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAATAATAGAGAAATTTAAACCAAA 720
Db TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAATAATAGAGAAATTTAAACCAAA 8719
Qy 721 GAAAGTTGAAATTTAGAGGTAAGGCTTAAGCTTGTGCTTTGAAACAGTGTCTGCT 780
Db GAAAGTTGAAATTTAGAGGTAAGGCTTAAGCTTGTGCTTTGAAACAGTGTCTGCT 8779
Qy 781 ACTGGGAAAAAGGCAAGCTTGGCTTCTCTAATTAATGATACAGGACTCTGTAATTCAT 840
Db ACTGGGAAAAAGGCAAGCTTGGCTTCTCTAATTAATGATACAGGACTCTGTAATTCAT 8839
Qy 841 ATTTTGCATGATGTAAGTAAGAAATGAAGCGGGTGAATGGACATGCCAGTAATCCC 900
Db ATTTTGCATGATGTAAGTAAGAAATGAAGCGGGTGAATGGACATGCCAGTAATCCC 8899
Qy 901 AGCACTCTGGGAGACTGAAGTGGGAGATCATCTTGAGCTCAGGAGTTCAAGACAGGCTG 960
Db AGCACTCTGGGAGACTGAAGTGGGAGATCATCTTGAGCTCAGGAGTTCAAGACAGGCTG 8959

Qy 961 GGCAACTAAAAATTAAAAAATAAAAAATACTAATAATGTTTTT 1001
Db 8960 GGCAACTAAAAATTAAAAAATAAAAAATACTAATAATGTTTTT 9000
RESULT 4
ADR72875
ID ADR72875 standard; DNA; 11570 BP.
XX ADR72875;
XX DT 02-DEC-2004 (first entry)
XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.
XX kallikrein 5; tumour marker; ovarian cancer;
KW epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;
KLK5; ds; gene.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 2221..11247
FT /*tag= b
FT /product= "Human ovarian cancer-related tumour marker
FT kallikrein 5 (hk5) protein"
FT exon 2221..2293
FT /*tag= a
FT /number= 1
FT intron 2294..4761
FT /*tag= c
FT /number= 1
FT exon 4762..5023
FT /*tag= d
FT intron 5024..5762
FT /*tag= e
FT exon 5763..6019
FT /*tag= f
FT /number= 3
FT intron 6020..6104
FT /*tag= g
FT /number= 3
FT exon 6105..6238
FT /*tag= h
FT /number= 4
FT exon 6239..11091
FT /*tag= i
FT /number= 4
FT exon 11092..11247
FT /*tag= j
FT /number= 5
W02004075713-A2.
10-SEP-2004.
26-FEB-2004; 2004WO-CA000281.
26-FEB-2003; 2003US-0450406P.
(MOUN) MOUNT SINAI HOSPITAL.
Diamandis EP;
WPI; 2004-661815/64.
P-PSDB; ADR72873.
Kallikrein markers detection method for detecting ovarian cancer in
patient, involves detecting kallikrein markers and Cal25 in sample
obtained from patient, and comparing detected amounts with standard

PT amounts.
PS Example 2; SEQ ID NO 5; 102pp; English.
XX
XX
The invention relates to a novel method for detecting a plurality of
CC kallikrein markers associated with ovarian cancer. The method comprises
CC obtaining a sample from a patient and detecting in the sample a plurality
CC of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
CC the kallikrein markers are selected from the group consisting of
CC kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
CC kallikrein 11. The detected amounts of the kallikrein markers are
CC compared with standard amounts. The method of the invention may be useful
CC for detecting kallikrein markers associated with ovarian cancer in a
CC patient and thus for detecting ovarian cancer, particularly epithelial
CC ovarian carcinoma. The current sequence is that of the human ovarian
CC cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention
CC which encodes a secreted serine protease and is located at chromosome
CC 19q13.4.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 2.9e-239;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGAGAAAGACATCTCGCCGGCGGTGGTGGCTCACACCTGTAAATCCCAACACTTT 60
DB 8000 AACAGAGAAAGACATCTCGCCGGCGGTGGTGGCTCACACCTGTAAATCCCAACACTTT 8059
QY 61 GGGAGGCGGAGCGGTGCATCTAGTCTAGGAGATGGAGACCATCTCTGCTGCACACGG 120
DB 8060 GGGAGGCGGAGCGGTGCATCTAGTCTAGGAGATGGAGACCATCTCTGCTGCACACGG 8119
QY 121 TGAACCCCTGTCTTACTATAAACAACAAAAAATAGCCGGCGGTGGTGGCGGCCCTGT 180
DB 8120 TGAACCCCTGTCTTACTATAAACAACAAAAAATAGCCGGCGGTGGTGGCGGCCCTGT 8179
QY 181 AGTCCAGCTACTCGGGAGGCTGAGGAGAGATGGCTGAAACCCGGAGGCGGAACTT 240
DB 8180 AGTCCAGCTACTCGGGAGGCTGAGGAGAGATGGCTGAAACCCGGAGGCGGAACTT 8239
QY 241 CGAGTGAGCGGAGGTTCACCACTGCATCTCAGCTGGGCAACAGTGAGACTCCGTCT 300
DB 8240 CGAGTGAGCGGAGGTTCACCACTGCATCTCAGCTGGGCAACAGTGAGACTCCGTCT 8299
QY 301 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGGTGGCAATTTAAAAAC 360
DB 8300 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGGTGGCAATTTAAAAAC 8359
QY 361 TATTAGCCTTCTGTAGGCAAGTTAGTATCTTCTTTTCCAGACCTCAAGGTGTTTTT 420
DB 8360 TATTAGCCTTCTGTAGGCAAGTTAGTATCTTCTTTTCCAGACCTCAAGGTGTTTTT 8419
QY 421 TTGTTTGTGTTTTCATACGGTGTGTGCTGGGTGGCCACTAAAGCTACAGCAAG 480
DB 8420 TTGTTTGTGTTTTCATACGGTGTGTGCTGGGTGGCCACTAAAGCTACAGCAAG 8479
QY 481 AAATAATAACAACACTACAACTACTAATAACCAATAGTATAAATAATAGCATCTGGCTA 540
DB 8480 AAATAATAACAACACTACAACTACTAATAACCAATAGTATAAATAATAGCATCTGGCTA 8539
QY 541 ATTGCTGCACTGTTTTTAAGTGGTTTGCATGCCCTCAGCTCATTTAACTTACCTGTT 600
DB 8540 ATTGCTGCACTGTTTTTAAGTGGTTTGCATGCCCTCAGCTCATTTAACTTACCTGTT 8599
QY 601 ATTATTGGCCCTATTTTTCAACACAGGAGCCCAAGGCTCAGACGTTAACTAACAGCCTC 660
DB 8600 ATTATTGGCCCTATTTTTCAACACAGGAGCCCAAGGCTCAGACGTTAACTAACAGCCTC 8659
QY 661 TCAAAAGAAACTCTCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTAACACCAA 720
DB 8660 TCAAAAGAAACTCTCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTAACACCAA 8719

QY 721 GAAAGTTGAAATTTAGAGGTACAGGCTAAGCTTGTGTTGTTGAAACAGTGTCTGCT 780
DB 8720 GAAAGTTGAAATTTAGAGGTACAGGCTAAGCTTGTGTTGTTGAAACAGTGTCTGCT 8779
QY 781 ACTGGGAAAAAGGCAAGTCTTGGCTTTCCCTAATAATTGATACACAGGACTCTGTAATTCAT 840
DB 8780 ACTGGGAAAAAGGCAAGTCTTGGCTTTCCCTAATAATTGATACACAGGACTCTGTAATTCAT 8839
QY 841 ATTTTGCATGCATTAAGTAAGAAATGAAGCCGGTGCAATGGCACATGCCAGTAATCCC 900
DB 8840 ATTTTGCATGCATTAAGTAAGAAATGAAGCCGGTGCAATGGCACATGCCAGTAATCCC 8899
QY 901 AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCCAGCTG 960
DB 8900 AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCCAGCTG 8959
QY 961 GGCAACTAAAAATTAATAAAAAATAAAAAATACATAATGTTTTT 1001
DB 8960 GGCAACTAAAAATTAATAAAAAATAAAAAATACATAATGTTTTT 9000
RESULT 5
AAK72959
ID AAK72959 standard; DNA; 31749 BP.
XX
AC AAK72959;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27771.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytosolic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.

PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-0232397P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	2000US-0236327P.	XX	WPI; 2001-483426/52.	
PR	29-SEP-2000;	2000US-0236367P.	DR		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	29-SEP-2000;	2000US-0236370P.	PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	02-OCT-2000;	2000US-0236802P.	XX	Disclosure; SEQ ID NO 27771; 3071pp + Sequence Listing; English.	
PR	02-OCT-2000;	2000US-0237037P.	PS		
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
PR	02-OCT-2000;	2000US-0237040P.	CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic	
PR	13-OCT-2000;	2000US-0239935P.	CC	activity, and can be used in gene therapy and vaccine production. (i)	
PR	13-OCT-2000;	2000US-0239937P.	CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
PR	20-OCT-2000;	2000US-0240960P.	CC	treatment of diseases associated with inappropriate (I) expression. For	
PR	20-OCT-2000;	2000US-0241221P.	CC	example, they may be used to treat disorders associated with decreased	
PR	20-OCT-2000;	2000US-0241785P.	CC	expression by rectifying mutations or deletions in a patient's genome	
PR	20-OCT-2000;	2000US-0241786P.	CC	that affect the activity of (I) by expressing inactive proteins or to	
PR	20-OCT-2000;	2000US-0241787P.	CC	supplement the patients own production of (I). Additionally, (I)	
PR	20-OCT-2000;	2000US-0241808P.	CC	polynucleotides may be used to produce the secreted (I), by inserting the	
PR	20-OCT-2000;	2000US-0241809P.	CC	nucleic acids into a host cell and culturing the cell to express the	
PR	20-OCT-2000;	2000US-0241826P.	CC	protein. (I) proteins and polynucleotides may be used to prevent,	
PR	01-NOV-2000;	2000US-0244617P.	CC	diagnose and treat immune/haematopoietic-related diseases, especially	
PR	08-NOV-2000;	2000US-0246474P.	CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
PR	08-NOV-2000;	2000US-0246475P.	CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
PR	08-NOV-2000;	2000US-0246476P.	CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169	
PR	08-NOV-2000;	2000US-0246477P.	CC	represent sequences used in the exemplification of the present invention	
PR	08-NOV-2000;	2000US-0246478P.	XX		
PR	08-NOV-2000;	2000US-0246523P.	SQ	Sequence 31749 BP; 9613 A; 6041 C; 6509 G; 9586 T; 0 U; 0 Other;	
PR	08-NOV-2000;	2000US-0246524P.			
PR	08-NOV-2000;	2000US-0246525P.			
PR	08-NOV-2000;	2000US-0246526P.			
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246609P.			
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			
PR	08-NOV-2000;	2000US-0246613P.			
<hr/>					
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-0232397P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	2000US-0236327P.	XX	WPI; 2001-483426/52.	
PR	29-SEP-2000;	2000US-0236367P.	DR		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	29-SEP-2000;	2000US-0236370P.	PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	02-OCT-2000;	2000US-0236802P.	XX	Disclosure; SEQ ID NO 27771; 3071pp + Sequence Listing; English.	
PR	02-OCT-2000;	2000US-0237037P.	PS		
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
PR	02-OCT-2000;	2000US-0237040P.	CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic	
PR	13-OCT-2000;	2000US-0239935P.	CC	activity, and can be used in gene therapy and vaccine production. (i)	
PR	13-OCT-2000;	2000US-0239937P.	CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
PR	20-OCT-2000;	2000US-0240960P.	CC	treatment of diseases associated with inappropriate (I) expression. For	
PR	20-OCT-2000;	2000US-0241221P.	CC	example, they may be used to treat disorders associated with decreased	
PR	20-OCT-2000;	2000US-0241785P.	CC	expression by rectifying mutations or deletions in a patient's genome	
PR	20-OCT-2000;	2000US-0241786P.	CC	that affect the activity of (I) by expressing inactive proteins or to	
PR	20-OCT-2000;	2000US-0241787P.	CC	supplement the patients own production of (I). Additionally, (I)	
PR	20-OCT-2000;	2000US-0241808P.	CC	polynucleotides may be used to produce the secreted (I), by inserting the	
PR	20-OCT-2000;	2000US-0241809P.	CC	nucleic acids into a host cell and culturing the cell to express the	
PR	20-OCT-2000;	2000US-0241826P.	CC	protein. (I) proteins and polynucleotides may be used to prevent,	
PR	01-NOV-2000;	2000US-0244617P.	CC	diagnose and treat immune/haematopoietic-related diseases, especially	
PR	08-NOV-2000;	2000US-0246474P.	CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
PR	08-NOV-2000;	2000US-0246475P.	CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
PR	08-NOV-2000;	2000US-0246476P.	CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169	
PR	08-NOV-2000;	2000US-0246477P.	CC	represent sequences used in the exemplification of the present invention	
PR	08-NOV-2000;	2000US-0246478P.	XX		
PR	08-NOV-2000;	2000US-0246523P.	SQ	Sequence 31749 BP; 9613 A; 6041 C; 6509 G; 9586 T; 0 U; 0 Other;	
PR	08-NOV-2000;	2000US-0246524P.			
PR	08-NOV-2000;	2000US-0246525P.			
PR	08-NOV-2000;	2000US-0246526P.			
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246609P.			
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			
PR	08-NOV-2000;	2000US-0246613P.			
<hr/>					
Qy	1	AACAGAGAAAGACATCTCGCGCGGCTGCTGCCTCACACCTGTATCCAAACACTTT	60		
Db	9594	AAAAGTTGAAGAAAAGATCGCGCGCGCGCTGCTGCCTCACACCTGTATCCAAACACTTT	9653		
Qy	61	GGGAGCGCGAGCGCGGTGGATCACTAGTCTCAGAGATGGAGACCATCTGCTGACACGG	120		

Db 9654 GGGAGCCCAAGCGGGTGGATCACAAGGTTCAGAGATCGAGACCATCTGGCTAACACGG 9713
Qy 121 TGAACCCCTGTCTCTACTAAAAACACAAAAATTAGCCGGGCGTGGTGGCAGCGCCTGT 180
Db 9714 TGAACCCCGCTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGCAGCGCCTGT 9773
Qy 181 AGTCCAGCTACTCCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGCGGAACCTT 240
Db 9774 AGTCCAGCTACTCCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGCGAGCTT 9833
Qy 241 GCAGTGAGCGAGGTTCACCACTGCCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCT 300
Db 9834 GCAGTGAGCGAGGTTCGCCACTGCCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCT 9893
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATCACA 333
Db 9894 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGA 9926

RESULT 6

AAC89888
ID AAC89888 standard; DNA; 78925 BP.

AC AAC89888;

DT 12-MAR-2001 (first entry)

DE Human FN gene.

Human; FN; fibronectin; prostate cancer; biallelic marker; diagnosis; ds.

Human sapiens.

WO200058509-A2.

05-OCT-2000.

28-MAR-2000; 2000WO-IB000431.

29-MAR-1999; 99US-0126780P.

(GEST) GENSET.

Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;

WPI; 2000-594647/56.

Fibronectin polynucleotide and polypeptide sequences, useful for determining the predisposition of individuals to cancer, such as prostate cancer.

Claim 1; Page 156-177; 208pp; English.

The present sequence is given in a specification relating to an isolated or purified recombinant polynucleotide comprising a contiguous span of at least 12 nucleotides of a fibronectin (FN) gene. The methods and sequences are useful for determining the predisposition of individuals to cancer such as prostate cancer and for the prognosis/detection of an eventual treatment response to therapeutic agents acting against prostate cancer. Biallelic markers allow association studies to be performed to identify genes involved in complex traits

Sequence 78925 BP; 23711 A; 15472 C; 15779 G; 23840 T; 0 U; 123 Other;

Query Match 28.0%; Score 280; DB 3; Length 78925;

Best Local Similarity 90.9%; Pred. No. 4.9e-59;

Matches 298; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 AACGAGAAGACATCTCGGCGGGTGGCTCACACCTGTAAATCCCAACACTTT 60

Db 11345 AAAAGTTGAAGAAAGATCGCCGGGCGGTGGCTCAGCCTGTATCCAGCACTTT 11404

Qy 61 GGGAGCCGAGCGGGTGGATCCTAGGTTCAGAGATCGAGACCATCTGGCTGACACGG 120
Db 11405 GGGAGCCCAAGCGGGTGGATCACAAGGTTCAGAGATCGAGACCATCTGGCTAACACGG 11464
Qy 121 TGAACCCCTGTCTCTACTAAAAACACAAAAATTAGCCGGGCGTGGTGGCAGCGCCTGT 180
Db 11465 TGAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGCAGCGCCTGT 11524
Qy 181 AGTCCAGCTACTCCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGCGGAACCTT 240
Db 11525 AGTCCAGCTACTCCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGCGAGCTT 11584
Qy 241 GCAGTGAGCGAGGTTCACCACTGCCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCT 300
Db 11585 GCAGTGAGCGAGGTTCGCCACTGCCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCT 11644
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAAGAAA 328
Db 11645 CAAAAAAGAAAAAGAAAAAGAAAAAGAAA 11672

RESULT 7

AAL57703

ID AAL57703 standard; DNA; 96599 BP.

XX AAL57703;

DT 16-OCT-2003 (first entry)

Human prolactin receptor (PRLR) genomic DNA sequence.

Human; cancer; carcinoma; breast cancer; CA gene;

carcinoma-associated gene; prolactin receptor; PRLR;

cytokine class-1 receptor superfamily; autoimmune disease;

carcinoma-associated protein; CAP; antisense therapy; gene therapy;

cytostatic activity; animal model; DNA vaccine; cancer of the bone;

cancer of the pharynx; cancer of the digestive system; skin cancer;

cancer of the respiratory system; mesothelioma; soft tissue cancer;

cancer of the oral cavity; cancer of the urinary system;

cancer of the endocrine system; multiple myeloma; leukaemia; lymphoma;

Hodgkin's lymphoma; non-Hodgkin's lymphoma; gene; ds.

Homo sapiens.

US2003064383-A1.

03-APR-2003.

20-MAR-2002; 2002US-00105948.

22-DEC-2000; 2000US-00747377.

02-MAR-2001; 2001US-00798586.

08-NOV-2001; 2001US-00052482.

(MORR/) MORRIS D W.

(ENG/) ENGELHARD E K.

Morris DW, Engelhard EK;

WPI; 2003-576379/54.

Novel carcinoma-associated gene for use in diagnosing and treating carcinomas, especially breast cancers or lymphomas, such as Hodgkin's and non-Hodgkin's lymphoma.

Claim 1; Page 37-51; 55pp; English.

This invention relates to novel sequences for use in diagnosis and treatment of cancer, especially carcinomas including breast cancer, as well as the use of novel compositions in screening methods. In particular, the invention comprises recombinant nucleic acids from a carcinoma-associated (CA) gene (such as the prolactin receptor, PRLR gene). PRLR is a member of the cytokine class-1 receptor superfamily,

Db 49344 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 49372

RESULT 11

ADL15049

ID ADL15049 standard; DNA; 77781 BP.

XX

AC ADL15049;

XX

XX

DT 06-MAY-2004 (first entry)

XX

XX Human melanoma associated MAGE-like DNA for cancer treatment.

DE

XX ds; gene; cytostatic; gene therapy; binding moiety; medicine; imaging;

XX diagnosis; prognosis; mantle cell lymphoma; cancer.

KW

XX Homo sapiens.

OS

XX WO2003068268-A2.

FN

XX 21-AUG-2003.

PD

XX 13-FEB-2003; 2003WO-EP001461.

XX

XX 14-FEB-2002; 2002GB-00003480.

PR

XX 29-JUN-2002; 2002GB-00015095.

XX

XX (BIOI-) BIOINVENT INT AB.

PA

XX Ek S, Borrebaeck CAK, Ehinger M;

PI

XX WPI; 2003-697496/66.

DR

XX P-PSDB; ADL15047.

DR

XX

XX New compound for treating, imaging, diagnosing or prognosing mantle cell

PT lymphoma, comprises a binding moiety (e.g. antibody) that binds to a

PT protein (e.g. human autotoxin polypeptide), and a further moiety (e.g.

PT nucleic acid).

XX

XX Disclosure; SEQ ID NO 61; 342pp; English.

PS

XX

XX The invention relates to a compound comprising a binding moiety which

CC selectively binds to a protein or polypeptide listed in the specification

CC (e.g. human autotoxin polypeptide or human CD24 signal transducer

CC polypeptide), and a further moiety. The compound is useful in medicine or

CC in the treatment, imaging, diagnosis or prognosis of mantle cell

CC lymphomas (MCL). It is used in preparing a medicament for treating MCL, a

CC diagnostic or prognostic agent for MCL, or an agent for imaging MCL cells

CC in the body of an individual. This sequence corresponds to a gene

CC encoding one of the polypeptides of the invention.

XX

SQ Sequence 77781 BP; 20855 A; 16833 C; 18350 G; 21743 T; 0 U; 0 Other;

Query Match 27.9%; Score 278.8; DB 10; Length 77781;

Best Local Similarity 89.1%; Pred. No. 9.7e-59;

Matches 301; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 AACGAGAAAGCACATCTCGCCGGCGTGGTGGCTCACACCTGTATCCCAACTTT 60

Db 60971 ATCTGGAATAAGAACACAGGGCGCGCGTGGCTCACGCCCTGTATCCCACTTT 61030

QY 61 GGGAGGCGGCGGTGGATCACTAGCTCAGGAGATGGAGACCATCTGGCTGACACGG 120

Db 61031 GGGAGGCGGCGGTGGATCAGAGTCAAGGATCAGACCATCTGGCTGACACGG 61090

QY 121 TGAACCCCTGTCTTACTATAAACAACAAAAAATTAGCCGGCGTGGTGGCAGGCGCTGT 180

Db 61091 TGAACCCCGTCTTACTATAAATAACAAAAAATTAGTGGCGTGGTGGCGGCGCTGT 61150

QY 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTCAACCCGGAGCGCACTT 240

Db 61151 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCGTGAACCCGGAGCGGAGCTT 61210

QY 241 GCACTGAGCCGAGGTTGCACCACTGCCTCAGCTGGGCAACACAGTGAAGACTCCGTC 300

Db 61211 GCACTGAGCCGAGATCGGCCACTGTCACTCAGCTGGGCGACAGAGCGAGACTCCGTC 61270

QY 301 CAAAAAAGAAAAAGAAAAAGAAAAAGAAATCACATCTCA 338

Db 61271 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAACACA 61308

RESULT 12

ADS36499

ID ADS36499 standard; DNA; 110300 BP.

XX

XX ADS36499;

XX

DT 16-DEC-2004 (first entry)

XX

XX Human autoimmune disease-related genomic DNA sequence - SEQ ID 1713.

DE

XX single nucleotide polymorphism detection; SNP detection;

XX rheumatoid arthritis; type 1 diabetes; multiple sclerosis;

XX systemic lupus erythematosus; inflammatory bowel disease; psoriasis;

KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;

KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;

KW primary systemic vasculitis; ds.

XX

XX Homo sapiens.

OS

XX WO2004083403-A2.

FN

XX 30-SEP-2004.

PD

XX 18-MAR-2004; 2004WO-US008461.

XX

XX 18-MAR-2003; 2003US-0455444P.

PR

XX 25-APR-2003; 2003US-0465241P.

XX

XX (APPL-) APPLERA CORP.

PA

XX Cargill M, Begovich AB, Alexander HC;

PI

XX WPI; 2004-728480/71.

DR

XX

XX New isolated nucleic acid molecule comprises at least 8 contiguous

PT nucleotides where one of the nucleotides is a single nucleotide

PT polymorphism (SNP), useful for diagnosing or treating autoimmune

PT diseases, e.g. rheumatoid arthritis.

XX

XX Claim 16; SEQ ID NO 1713; 123pp; English.

PS

XX

XX The invention comprises amino acid and coding sequences containing

CC genetic polymorphisms associated with an altered risk of developing an

CC autoimmune disease (e.g. rheumatoid arthritis). The invention further

CC comprises a method of identifying an individual that has an altered risk

CC of developing an autoimmune disease, comprising detecting a single

CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA

CC and protein sequences of the invention are useful for diagnosing and

CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1

CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory

CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious

CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,

CC myocardiitis, Sjogren's disease, or primary systemic vasculitis. The

CC present nucleic acid represents a human autoimmune disease-related

CC genomic DNA sequence of the invention. NOTE: The present sequence is not

CC shown in the specification, but has been retrieved from the WIPO website.

XX

SQ Sequence 110300 BP; 33141 A; 18377 C; 20462 G; 38206 T; 0 U; 114 Other;

Query Match 27.9%; Score 278.8; DB 13; Length 110300;

Best Local Similarity 86.7%; Pred. No. 1.1e-58;

Matches 307; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 11 AGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTTGGAGCGCGA 70

```
Db 29024 AGTCCATCTCGCGCGCGGTGGCTCAGCCCTGTANTCCAGCACTTTGGGAGGCCGA 29083
Qy 71 GCGGGTGGATCACTAGTGTAGGAGATGGAGACCATCTTGCTGACACGGTGAAACCCCTG 130
Db 29084 GCGCGCGCGATCACAGGTGAGGATCGAGACCATCTTGCTGCTAACACGGTGAAACCCCG 29143
Qy 131 TCTCTACTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCCTGTAGTCCAGCT 190
Db 29144 TCTCTACTAAAAAATACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCCTGTAGTCCAGCT 29203
Qy 191 ACTCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGGCGGAACTTGCAGTGAACC 250
Db 29204 ACTCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGGCGGAGCTTGCAGTGAACC 29263
Qy 251 GAGTTGACCACTGCACTCCAGCTTGGCGAACACAGTGAGACTCCGCTTCAAAAAAAA 310
Db 29264 GAGATCGCGCCACTGCACTCCAGCTTGGCGAACACAGTGAGACTCCGCTTCAAAAAAAA 29323
Qy 311 AAAAGAAAGAAAGAAATACATCTCATTCAGTGGTGGCATTTAAACTATT 364
Db 29324 AAAAGAAAGAAAGTAGTCCATCTCTCTGTCTGAAAGAGATTATAGAAATT 29377

RESULT 13
AAK90223
ID AAK90223 standard; DNA; 3407 BP.
XX AC AAK90223;
XX DT 05-NOV-2001 (first entry)
XX DE Human digestive system antigen genomic sequence SEQ ID NO: 3799.
XX KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX KW digestive system disorder; Meckel's diverticulum; ds.
XX OS Homo sapiens.
XX FN WO200155314-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001324.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241878P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
```


Qy	6	AAGAAAGC	A	C	T	C	G	C	G	G	C	T	G	T	G	G	C	T	A	C	A	C	T	T	T	G	G	A	G	65																
Db	3068	AAAAAGTG	A	A	A	A	A	G	T	G	A	A	C	A	C	G	C	G	G	C	G	T	G	C	T	A	C	A	C	T	T	T	G	G	A	G	3127									
Qy	66	GCCGAGCG	C	G	G	T	G	G	A	T	C	A	C	T	A	G	T	C	A	G	A	G	A	T	G	A	G	A	G	A	G	A	G	A	G	125										
Db	3128	GCCGAGCG	C	G	G	T	G	G	A	T	C	A	C	A	G	T	C	A	G	A	G	A	T	G	A	G	A	G	A	G	A	G	A	G	3187											
Ov	126	CCCTGTCT	C	T	C	T	A	A	A	A	A	C	A	A	A	A	A	T	T	A	G	C	G	G	G	C	G	T	G	T	G	C	A	G	C	G	C	T	G	T	A	G	T	C	C	185

Db	3188	CCCCGCTCTCTACTAAAAATACAAAAATAGCCGGCGGTGGCCGGCGCCCTGTAGTCC	3247	PR 01-SEP-2000; 2000US-0222877P
Qy	186	CAGCTACTCCGGAGGCTGAGGCAGGAATGGCTGAACCCGGGAGCGGAACTTCGAGT	245	PR 01-SEP-2000; 2000US-0229343P
Db	3248	CAGCTACTCCGGAGGCTGAGGCAGGAATGGCGTGAACCCGGGAGCGGAGCTTCGAGT	3307	PR 01-SEP-2000; 2000US-0229344P
Qy	246	GAGCCGAGTTGCAACCACTGCACCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCTCAAAA	305	PR 01-SEP-2000; 2000US-0229509P
Db	3308	GAGCCGAGATCGCGCCACTGCACCTCCAGCCTGGGCGAGCAGCGAGACTCCGCTCTCAAAA	3367	PR 05-SEP-2000; 2000US-0229513P
Qy	306	AAAAAAAAAGAAAGAAAGAAA 328		PR 06-SEP-2000; 2000US-0230437P
Db	3368	AAAAAAAAAAAAAAAAAAAAAA 3390		PR 06-SEP-2000; 2000US-0230438P
RESULT 15				PR 08-SEP-2000; 2000US-0231242P
ADB32812				PR 08-SEP-2000; 2000US-0231244P
ID	ADB32812	standard; DNA; 3407 BP.		PR 08-SEP-2000; 2000US-0231413P
XX	AC	ADB32812;		PR 08-SEP-2000; 2000US-0232080P
XX	AD	04-DEC-2003 (first entry)		PR 08-SEP-2000; 2000US-0232081P
XX	DE	Human novel colon related polypeptide DNA SEQ ID NO 749.		PR 12-SEP-2000; 2000US-0231968P
KW	KW	gene therapy; cancer; liver disorder; hepatitis; neural disorder;		PR 14-SEP-2000; 2000US-0233977P
KW	KW	Alzheimer's disease; human; colon; ds.		PR 14-SEP-2000; 2000US-0233987P
XX	OS	Homo sapiens.		PR 14-SEP-2000; 2000US-0234000P
XX	PN	US2003050231-A1.		PR 14-SEP-2000; 2000US-0234013P
XX	PD	13-MAR-2003.		PR 14-SEP-2000; 2000US-0233063P
XX	PF	17-JAN-2001; 2001US-00764872.		PR 14-SEP-2000; 2000US-0233064P
XX	PR	31-FEB-2000; 2000US-0179065P.		PR 14-SEP-2000; 2000US-0233065P
PR	PR	04-FEB-2000; 2000US-0180628P.		PR 21-SEP-2000; 2000US-0234223P
PR	PR	24-FEB-2000; 2000US-0184664P.		PR 21-SEP-2000; 2000US-0234277P
PR	PR	02-MAR-2000; 2000US-0186350P.		PR 25-SEP-2000; 2000US-0234997P
PR	PR	16-MAR-2000; 2000US-0189874P.		PR 25-SEP-2000; 2000US-0234998P
PR	PR	17-MAR-2000; 2000US-0190076P.		PR 26-SEP-2000; 2000US-0234844P
PR	PR	18-APR-2000; 2000US-0198123P.		PR 27-SEP-2000; 2000US-0235834P
PR	PR	19-MAY-2000; 2000US-0205515P.		PR 29-SEP-2000; 2000US-0236327P
PR	PR	07-JUN-2000; 2000US-0209467P.		PR 29-SEP-2000; 2000US-0236367P
PR	PR	28-JUN-2000; 2000US-0214886P.		PR 29-SEP-2000; 2000US-0236369P
PR	PR	30-JUN-2000; 2000US-0215135P.		PR 29-SEP-2000; 2000US-0236370P
PR	PR	07-JUL-2000; 2000US-0216647P.		PR 02-OCT-2000; 2000US-0236802P
PR	PR	11-JUL-2000; 2000US-0217487P.		PR 02-OCT-2000; 2000US-0237037P
PR	PR	11-JUL-2000; 2000US-0217496P.		PR 02-OCT-2000; 2000US-0237038P
PR	PR	14-JUL-2000; 2000US-0218290P.		PR 02-OCT-2000; 2000US-0237039P
PR	PR	26-JUL-2000; 2000US-0220963P.		PR 13-OCT-2000; 2000US-0239335P
PR	PR	26-JUL-2000; 2000US-0220964P.		PR 13-OCT-2000; 2000US-0239337P
PR	PR	14-AUG-2000; 2000US-0224518P.		PR 20-OCT-2000; 2000US-0240960P
PR	PR	14-AUG-2000; 2000US-0224519P.		PR 20-OCT-2000; 2000US-0241221P
PR	PR	14-AUG-2000; 2000US-0225213P.		PR 20-OCT-2000; 2000US-0241785P
PR	PR	14-AUG-2000; 2000US-0225214P.		PR 20-OCT-2000; 2000US-0241786P
PR	PR	14-AUG-2000; 2000US-0225266P.		PR 20-OCT-2000; 2000US-0241787P
PR	PR	14-AUG-2000; 2000US-0225267P.		PR 20-OCT-2000; 2000US-0241808P
PR	PR	14-AUG-2000; 2000US-0225268P.		PR 20-OCT-2000; 2000US-0241809P
PR	PR	14-AUG-2000; 2000US-0225269P.		PR 20-OCT-2000; 2000US-0241826P
PR	PR	14-AUG-2000; 2000US-0225447P.		PR 01-NOV-2000; 2000US-0244617P
PR	PR	14-AUG-2000; 2000US-0225477P.		PR 08-NOV-2000; 2000US-0246474P
PR	PR	14-AUG-2000; 2000US-0225757P.		PR 08-NOV-2000; 2000US-0246475P
PR	PR	14-AUG-2000; 2000US-0225758P.		PR 08-NOV-2000; 2000US-0246476P
PR	PR	14-AUG-2000; 2000US-0225759P.		PR 08-NOV-2000; 2000US-0246477P
PR	PR	18-AUG-2000; 2000US-0225759P.		PR 08-NOV-2000; 2000US-0246478P
PR	PR	22-AUG-2000; 2000US-0226279P.		PR 08-NOV-2000; 2000US-0246523P
PR	PR	22-AUG-2000; 2000US-0226681P.		

PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249222P.
PR 17-NOV-2000; 2000US-0249223P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249225P.
PR 17-NOV-2000; 2000US-0249226P.
PR 17-NOV-2000; 2000US-0249227P.
PR 17-NOV-2000; 2000US-0249228P.
PR 17-NOV-2000; 2000US-0249229P.
PR 17-NOV-2000; 2000US-0249230P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-625420/59.

New nucleic acid molecule, useful for preparing a medicament for preventing, treating or ameliorating a medical condition e.g. cancer, liver disorders or neural disorders.

Disclosure; SEQ ID NO 749; 216pp; English.

The invention relates to an isolated nucleic acid molecule encoding a polypeptide. The nucleic acid is useful for preparing a medicament for preventing, treating or ameliorating a medical condition e.g. cancer, liver disorders such as hepatitis or neural disorders such as Alzheimer's disease. The present sequence represents a human novel colon related polypeptide DNA. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030050231.

Sequence 3407 BP; 1134 A; 634 C; 666 G; 973 T; 0 U; 0 Other;

Query Match 27.8%; Score 278.2; DB 9; Length 3407;
Best Local Similarity 91.3%; Pred. No. 4.6e-59;
Matches 295; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 6 AAGAAAGCACATCTCGCGCGCGGTGGTGTACACCTGTATCCCAACACTTTGGGAG 65
DB 3068 AAAAAAGTGAACACGGCGCGGTGGTGTACACCTGTATCCCAACACTTTGGGAG 3127
QY 66 GCCGAGGCGGTGATCTAGTTCAGGAGATGGAGACATCTTGGCTGACACGGTGAAA 125
DB 3128 GCCGAGGCGGTGATCTAGTTCAGGAGATGGAGACATCTTGGCTGACACGGTGAAA 3187
QY 126 CCCTGTCTCTACTAAACACAAAAAATAGCCGGCGGTGGTGGCAGCGCTGTAGTCC 185
DB 3188 CCCCGTCTCTACTAAACACAAAAAATAGCCGGCGGTGGTGGCAGCGCTGTAGTCC 3247
QY 186 CAGTACTCTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGGAGGCGGAACCTTGCACT 245
DB 3248 CAGTACTCTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGGAGGCGGAACCTTGCACT 3307

QY 246 GAGCCGAGTTTGACCACTGCACCTCCAGCCTGGGCAACACAGTGCAGCTCCGCTCTCAAAA 305
DB 3308 GAGCCGAGATCGCCCACTGCACCTCCAGCCTGGGCAACACAGTGCAGCTCCGCTCTCAAAA 3367
QY 306 AAAAAAAGAAAAAGAAAAAGAAA 328
DB 3368 AAAAAAAGAAAAAGAAAAAGAAA 3390

Search completed: July 4, 2005, 03:03:31
Job time : 358.271 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:41:10 ; Search time 113.609 seconds

(without alignments)
14417.061 Million cell updates/sec

Title: US-09-936-271C-13_COPY_8000_9000

Perfect score: 1001

Sequence: 1 aacagaagaagcaccatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	10818	4	US-09-949-016-13583
c 2	600.6	60.0	601	4	US-09-949-016-13583
c 3	600.6	60.0	601	4	US-09-949-016-13583
c 4	599.4	59.9	601	4	US-09-949-016-13583
c 5	533.6	53.3	601	4	US-09-949-016-13583
c 6	518.6	51.8	601	4	US-09-949-016-13583
c 7	470.6	47.0	601	4	US-09-949-016-13583
c 8	465.6	46.5	601	4	US-09-949-016-13583
c 9	280.2	28.0	84252	4	US-09-949-016-17315
c 10	279.4	27.9	169334	4	US-09-949-016-15999
c 11	278.4	27.8	38052	4	US-09-949-016-13825
c 12	277.6	27.7	130724	4	US-09-949-016-13753
c 13	277.4	27.7	108341	4	US-09-949-016-16621
c 14	277.4	27.7	115814	4	US-09-949-016-16205
c 15	276.8	27.7	601	4	US-09-949-016-130842
c 16	276.8	27.7	20495	4	US-09-949-016-17198
c 17	276.8	27.7	114793	4	US-10-148-806-3
c 18	276.8	27.7	134292	4	US-09-949-016-12158
c 19	276.6	27.6	121982	4	US-09-949-016-12085
c 20	276.6	27.6	121982	4	US-09-949-016-14105
c 21	276.4	27.6	181429	4	US-09-949-016-12372
c 22	276.4	27.6	181430	4	US-09-949-016-15772
c 23	276.2	27.6	84425	4	US-09-949-016-17402
c 24	275.2	27.5	133719	4	US-09-949-016-15092
c 25	275	27.5	68778	4	US-09-949-016-16406
c 26	274.6	27.4	23766	4	US-09-949-016-13569
c 27	274.4	27.4	636591	4	US-09-949-016-11808

c 28	274.4	27.4	636591	4	US-09-949-016-13388	Sequence 13388, A
c 29	274.2	27.4	14543	4	US-09-959-873B-17	Sequence 17, Appl
c 30	273.4	27.3	3243	4	US-03-949-016-12186	Sequence 12186, A
c 31	273.4	27.3	60465	4	US-09-949-016-15995	Sequence 15995, A
c 32	273.4	27.3	76264	4	US-09-949-016-15773	Sequence 15773, A
c 33	273.2	27.3	14636	3	US-09-173-914-6	Sequence 6, Appli
c 34	273	27.3	192956	4	US-09-949-016-14382	Sequence 14382, A
c 35	272.8	27.3	154023	4	US-09-949-016-17057	Sequence 17057, A
c 36	272.6	27.2	15392	4	US-09-949-016-13596	Sequence 13596, A
c 37	272.6	27.2	48316	4	US-09-949-016-12853	Sequence 12853, A
c 38	272.6	27.2	48316	4	US-09-949-016-13405	Sequence 13405, A
c 39	272.6	27.2	104428	4	US-09-949-016-12737	Sequence 12737, A
c 40	272.6	27.2	104429	4	US-09-949-016-13814	Sequence 13814, A
c 41	272.4	27.2	33379	4	US-09-949-016-13861	Sequence 13861, A
c 42	272.4	27.2	118382	4	US-09-949-016-15996	Sequence 15996, A
c 43	272.4	27.2	118382	4	US-09-949-016-15997	Sequence 15997, A
c 44	272.2	27.2	601	4	US-09-949-016-22105	Sequence 22105, A
c 45	272.2	27.2	601	4	US-09-949-016-154105	Sequence 154105, A

ALIGNMENTS

RESULT 1

US-09-949-016-13583
; Sequence 13583, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13583
; LENGTH: 10818
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13583

Query Match	100.0%	Score	1001;	DB	4;	Length	10818;
Best Local Similarity	100.0%;	Pred. No.	6.5e-257;	Mismatches	0;	Indels	0;
Matches	1001;	Conservative	0;				
Qy	1	AACAGAGAAAGACATCTCGCGCGGGGTGGTGGCTCACACCTGTATCCACACATTT	60				
Db	5245	AACAGAGAAAGACATCTCGCGCGGGGTGGTGGCTCACACCTGTATCCACACATTT	5304				
Qy	61	GGGAGCGCGAGCGGGTGGATCACTAGGTACAGAGATGAGACCATCTCGCTGACACGG	120				
Db	5305	GGGAGCGCGAGCGGGTGGATCACTAGGTACAGAGATGAGACCATCTCGCTGACACGG	5364				
Qy	121	TGAACCCCTGTCTCTACTATAAAACACAAAAATAGCCGGCGGTGGTGGCGCCCTGT	180				
Db	5365	TGAACCCCTGTCTCTACTATAAAACACAAAAATAGCCGGCGGTGGTGGCGCCCTGT	5424				
Qy	181	AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGGAGCGGAATTT	240				
Db	5425	AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGGAGCGGAATTT	5484				
Qy	241	GCAGTGAGCCGAGGTTCACACCTGCATCTCAGCGCTGGGCAACACAGTGAGACTCCGCT	300				
Db	5485	GCAGTGAGCCGAGGTTCACACCTGCATCTCAGCGCTGGGCAACACAGTGAGACTCCGCT	5544				
Qy	301	CAAAAAAAGAAAAAGAAAAAATCATCTCATTCAGTGGTGGCATTTAAAAAC	360				

Query Match	60.0%;	Score	600.6;	DB 4;	Length	601;			
Best Local Similarity	99.8%;	Pred. No.	1.3e-150;						
Matches	600;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	171	AGGCGCCTGTAGTCCCAAGCTACTCGGAGGCGCTGAGGCAGGAGAGATGGCTCGAACC	GGGA	230					
DB	601	AGGCGCCTGTAGTCCCAAGCTACTCGGAGGCGCTGAGGCAGGAGAGATGGCTCGAACC	GGGA	542					
QY	231	GGCGGAATTCGCTAGTGACCGAGGTTCACCACTGCACCTCCAGGCTGGGCAACACAGTGA	290						
DB	541	GGCGGAATTCGCTAGTGACCGAGGTTCACCACTGCACCTCCAGGCTGGGCAACACAGTGA	482						
QY	291	GACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGTGG	350						
DB	481	GACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGTGG	422						
QY	351	CAATTTAAAACTATTTAGCCCTTCTGTAGGCAAGTTAGTATCTTGTGTTTTTCCAGACCTCA	410						
DB	421	CAATTTAAAACTATTTAGCCCTTCTGTAGGCAAGTTAGTATCTTGTGTTTTTCCAGACCTCA	362						
QY	411	AGTGTTTTTTTTGTTTTGTTTTTTCATACCGGTGTGTGTCTGGGTGGGCCACTAAAAAGC	470						
DB	361	AGTGTTTTTTTTGTTTTGTTTTTTCATACCGGTGTGTGTCTGGGTGGGCCACTAAAAAGC	302						
QY	471	TACAGCAAGAAATAATAACAACTACAACTACTAATACCAATAGTATAAAAAATAATAG	530						
DB	301	TACAGCAAGAAATAATAACAACTACTAATACCAATAGTATAAAAAATAATAG	242						
QY	531	CATCTGCTTAATTCGCTGCACACTGTGTTTAAAGTGGTTTGCATGCTCAGCTCATTAACCTCA	590						
DB	241	CATCTGCTTAATTCGCTGCACACTGTGTTTAAAGTGGTTTGCATGCTCAGCTCATTAACCTCA	182						
QY	591	TTTACCTGTTATTATTGCGCCCTATTTTACAAACAAGGAGGCAAGGCTCAGAGCAGTTAAAC	650						
DB	181	TTTACCTGTTATTATTGCGCCCTATTTTACAAACAAGGAGGCAAGGCTCAGAGCAGTTAAAC	122						
QY	651	TAAACAGCCTCTCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAATAATAGAGAGAAT	710						
DB	121	TAAACAGCCTCTCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAATAATAGAGAGAAT	62						
QY	711	TAAACCAACAAGAAAGTTGAAATTTAGAGGTACAGGCAGCTAAGCTTGTTGCTTTTGAAC	770						
DB	61	TAAACCAACAAGAAAGTTGAAATTTAGAGGTACAGGCAGCTAAGCTTGTTGCTTTTGAAC	2						
QY	771	A	771						
DB	1	A	1						

RESULT 4
US-09-949-016-63410/c
; Sequence 63410, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63410
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-63410					
Query Match 59.9%; Score 599.4; DB 4; Length 601;					
Best Local Similarity 99.8%; Pred. No. 2.6e-150;					
Matches 600; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
Qy	10	AAGCACATCTCGGCGGGCGGTGGCTACACTGTAAATCCCAACACTTTGGAGGCCG	69		
Dd	601	AAGCACATCTCGGCGGGCGGTGGCTACACTGTAAATCCCAACACTTTGGAGGCCG	542		
Qy	70	AGGCGGTGGATCACTAGGTCAGAGATGGAGACCATTCTTGGCTGACACGGTGAAAACCCCT	129		
Dd	541	AGGCGGTGGATCACTAGGTCAGAGATGGAGACCATTCTTGGCTGACACGGTGAAAACCCCT	482		
Qy	130	GTCTCTACTAAAAACAACAAAAATTAGCGGGCGGTGGCGAGCGCCTGTAGTCCCAGC	189		
Dd	481	GTCTCTACTAAAAACAACAAAAATTAGCGGGCGGTGGCGAGCGCCTGTAGTCCCAGC	422		
Qy	190	TACTCGGGAGGCTGAGGCAGGAATAATGGCTGAAACCGGGAGCGGAACCTTGCAGTGAGC	249		
Dd	421	TACTCGGGAGGCTGAGGCAGGAATAATGGCTGAAACCGGGAGCGGAACCTTGCAGTGAGC	362		
Qy	250	CGAGGTTGCACCACTGCACCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCAAAAAAAA	309		
Dd	361	CGAGGTTGCACCACTGCACCTCCAGCCTGGGCAACACAGTGAGACTCCGCTCAAAAAAAA	302		
Qy	310	AAAAAGAAAAGAAAAGAAATCACATCTCAATCCAAGTGGTGGCATTTAAAACTATTTAGCC	369		
Dd	301	CAAAGAAGAAAAGAAAAGAAATCACATCTCAATCCAAGTGGTGGCATTTAAAACTATTTAGCC	242		
Qy	370	TTTTCTGTAGCAAGTTAGTAATCTGTTTTTCAGACCTCAAGGTGTTTTTTTGTGTGT	429		
Dd	241	TTTTCTGTAGCAAGTTAGTAATCTGTTTTTCAGACCTCAAGGTGTTTTTTTGTGTGT	182		
Qy	430	TTTTTCATACCGGTGTGTGTTCTGGGTGTGGCCACTAAAGCTACAAGCAAGAATAATAA	489		
Dd	181	TTTTTCATACCGGTGTGTGTTCTGGGTGTGGCCACTAAAGCTACAAGCAAGAATAATAA	122		
Qy	490	CAACTACAACTACTAAATAACCAATAGTATAAAAATAATAGCATCTGGCTAATTGCTGGA	549		
Dd	121	CAACTACAACTACTAAATAACCAATAGTATAAAAATAATAGCATCTGGCTAATTGCTGGA	62		
Qy	550	CACGTGTTTAAAGTGGTTTGCATGCCCTCAGCTCAATTAACCTATTTACTGTATATTGGC	609		
Dd	61	CACGTGTTTAAAGTGGTTTGCATGCCCTCAGCTCAATTAACCTATTTACTGTATATTGGC	2		
Qy	610	C 610			
Dd					
Qy		1 C 1			
RESULT 5					
US-09-949-016-63409/c					
; Sequence 63409, Application US/09949016					
; Patent No. 6812339					
; GENERAL INFORMATION:					
; APPLICANT: VENTER, J. Craig et al.					
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED					
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF					
; FILE REFERENCE: CL001307					
; CURRENT APPLICATION NUMBER: US/09/949,016					
; CURRENT FILING DATE: 2000-04-14					
; PRIOR APPLICATION NUMBER: 60/241,755					
; PRIOR FILING DATE: 2000-10-20					
; PRIOR APPLICATION NUMBER: 60/237,768					
; PRIOR FILING DATE: 2000-10-03					
; PRIOR APPLICATION NUMBER: 60/231,498					
; PRIOR FILING DATE: 2000-09-08					
; NUMBER OF SEQ ID NOS: 207012					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 63409					
; LENGTH: 601					
; TYPE: DNA					

```

RESULT 5
US-09-949-016-63409/c
; Sequence 63409, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63409
; LENGTH: 601
; TYPE: DNA

```

; ORGANISM: Human
US-09-949-016-63409

Query Match 53.3%; Score 533.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 9.6e-133; Mismatches 0; Indels 0; Gaps 0;
Matches 533; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 60
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 460

Qy 61 GGGAGCGCGAGCGGTGGATCACTAGTCAAGGAGATGGAGACCATCTGCTGCACACGG 120
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 475

Qy 61 GGGAGCGCGAGCGGTGGATCACTAGTCAAGGAGATGGAGACCATCTGCTGCACACGG 120
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 475

Qy 474 GGGAGCGCGAGCGGTGGATCACTAGTCAAGGAGATGGAGACCATCTGCTGCACACGG 415
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 475

Qy 121 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCCGGCGGTGGTGGAGCGCCCTGT 180
Db TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCCGGCGGTGGTGGAGCGCCCTGT 180

Qy 414 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCCGGCGGTGGTGGAGCGCCCTGT 355
Db TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCCGGCGGTGGTGGAGCGCCCTGT 355

Qy 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240
Db AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240

Qy 354 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCCGGAGCGGAACTT 295
Db AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCCGGAGCGGAACTT 295

Qy 241 GCAGTGAGCGGAGGTGGACACTGCACTCAGCCTGGGGAACACAGTGAACCTCCGTCT 300
Db GCAGTGAGCGGAGGTGGACACTGCACTCAGCCTGGGGAACACAGTGAACCTCCGTCT 300

Qy 294 GCAGTGAGCGGAGGTGGACACTGCACTCAGCCTGGGGAACACAGTGAACCTCCGTCT 235
Db GCAGTGAGCGGAGGTGGACACTGCACTCAGCCTGGGGAACACAGTGAACCTCCGTCT 235

Qy 301 CAAAAAAGAAAAAGAAAAAGAAATACATCTCAATCAAGTGGTGGCATTTAAAC 360
Db CAAAAAAGAAAAAGAAAAAGAAATACATCTCAATCAAGTGGTGGCATTTAAAC 175

Qy 361 TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGGTGTTTTT 420
Db TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGGTGTTTTT 115

Qy 421 TTGTTTGTGTTTTCATACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480
Db TTGTTTGTGTTTTCATACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480

Qy 114 TTGTTTGTGTTTTCATACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 55
Db TTGTTTGTGTTTTCATACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 55

Qy 481 AATAATAAACAACTACAACTACTAATACCAATAGTATATAAATAATAGCATC 534
Db AATAATAAACAACTACAACTACTAATACCAATAGTATATAAATAATAGCATC 1

RESULT 6

US-09-949-016-63408/c
; Sequence 63408, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63408
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63408

Query Match 51.8%; Score 518.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 9.7e-125; Mismatches 0; Indels 0; Gaps 0;
Matches 518; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 60
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 460

Qy 61 GGGAGCGCGAGCGGTGGATCACTAGTCAAGGAGATGGAGACCATCTGCTGCACACGG 120
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 459

Qy 121 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCCGGCGGTGGTGGAGCGCCCTGT 180
Db TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCCGGCGGTGGTGGAGCGCCCTGT 340

Qy 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240
Db AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCCGGAGCGGAACTT 280

Qy 241 GCAGTGAGCGGAGGTGGACACTGCACTCAGCCTGGGGAACACAGTGAACCTCCGTCT 300
Db GCAGTGAGCGGAGGTGGACACTGCACTCAGCCTGGGGAACACAGTGAACCTCCGTCT 220

Qy 301 CAAAAAAGAAAAAGAAAAAGAAATACATCTCAATCAAGTGGTGGCATTTAAAC 360
Db CAAAAAAGAAAAAGAAAAAGAAATACATCTCAATCAAGTGGTGGCATTTAAAC 160

Qy 361 TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGGTGTTTTT 420
Db TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGGTGTTTTT 100

Qy 421 TTGTTTGTGTTTTCATACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480
Db TTGTTTGTGTTTTCATACCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 40

Qy 481 AATAATAAACAACTACAACTACTAATACCAATAGTAT 519
Db AATAATAAACAACTACAACTACTAATACCAATAGTAT 1

RESULT 7

US-09-949-016-63407/c
; Sequence 63407, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63407
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63407

Query Match 47.0%; Score 470.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 6.3e-116; Mismatches 1; Indels 0; Gaps 0;
Matches 470; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 60
Db AACAGAGAAAGCACATCTCGCGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 412

Qy 61 GGGAGCGCGAGCGGTGGATCACTAGTCAAGGAGATGGAGACCATCTGCTGCACACGG 120
Db GGGAGCGCGAGCGGTGGATCACTAGTCAAGGAGATGGAGACCATCTGCTGCACACGG 352

GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15999
; LENGTH: 169334
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(169334)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15999

Query Match 27.9%; Score 279.4; DB 4; Length 169334;
Best Local Similarity 90.6%; Pred. No. 9.1e-64;
Matches 298; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 5 GAAGAAAGACATCTCGGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTTGGGA 64
Db 41044 GAAATAACATCTGGCGGGCGGTGGCTCACGCTGTATATCCAGCACTTTGGGA 41103

Qy 65 GGCGGAGCGGTGGATGACTAGTTCAGGAGATGGAGACCATCTGGCTGACACGGTGAA 124
Db 41104 GGCGGAGCGGGCGATCAGAGTTCAGGAGATCGAGACCATCTCCGGCTAAACCGTGAA 41163

Qy 125 ACCCTGTCTCTACTTAAACCAACAAAAATTAGCGGGCGTGGTGGCTGTAGTC 184
Db 41164 ACCCGTCTCTACTTAAATAACAAAAATTAGCGGGCGTGGTGGCTGTAGTC 41223

Qy 185 CCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTGAAACCGGAGCGGAACTTGCAG 244
Db 41224 CCAGCTACTTGGAGGCTGAGCGAGGAGATGGCTGAAACCGGAGCGGAGCTTGCAG 41283

Qy 245 TGAGCCGAGGTTGCACCACTGCACCTCCAGCTGGCGGAGGAGTTCCTCAAA 304
Db 41284 TGAGCCGAGATTGGCCCACTGCACCTCCAGCTGGCGGAGGAGACTCCGCTCAAA 41343

Qy 305 AAAAAAAGAAAGAAAGAAATTAAGCGGGCGTGGTGGCGAGCGGCTGTAGTC 184
Db 41344 AAAAAAAGAAAGAAAGAAATTAAGCGGGCGTGGTGGCGAGCGGCTGTAGTC 41372

RESULT 11
US-09-949-016-13825
; Sequence 13825, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
;

; SEQ ID NO 13825
; LENGTH: 38052
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(38052)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13825

Query Match 27.8%; Score 278.4; DB 4; Length 38052;
Best Local Similarity 58.5%; Pred. No. 8.5e-64;
Matches 566; Conservative 0; Mismatches 386; Indels 16; Gaps 4;

Qy 5 GAAGAAAGACATCTCGGCGGGCGTGGTGGCTCACACCTGTATATCCCAACACTTTGGGA 64
Db 3958 GAGTAAAAAATTTCAAGGCCGGGGCGGTGGCTCACGCTGTATATCCAGAACTTTGGGA 4017

Qy 65 GGCCGAGGCGGTGGATCACTAGGTCAAGAGATGGAGACCATCTCTGGCTGACACGGTGAA 124
Db 4018 GGCCGAGGCGGCAGATCACAGGTCAAGAGATCGATACAATCTCTGGCTAACACGGTGAA 4077

Qy 125 ACCCTGTCTCTACTTAAAAACACAAAAAATTAGCGGGCGTGGTGGCGGCGCTGTAGTC 184
Db 4078 ACCCGTCTCTACTTAAAAAATACAAAAAGTTAGCTGGGCGTGGTGGCGGCGCTGTAGTC 4137

Qy 185 CCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAAACCGGAGGCGGAACTTGCAG 244
Db 4138 CCAGCTACTCGAGAGGTTGAGGCGAGGAGATGGGCTGGACCCGGGAGCGGAGTTGCAG 4197

Qy 245 TGAGCCGAGGTTGCACCACTGCACCTCCAGCTGGGCAACACAGTGAGACTCCGCTCAAA 304
Db 4198 TGAGCCAAAGATCGCGCACTGCCTCCAGCTGGGCGACAGACAGACTCCGCTCAAA 4257

Qy 305 AAAAAAAGAAAGAAAGAAATTAAGCTTCAATCAAGTGGTGGCACTTAAACATTT 364
Db 4258 AAAAAAAGAAAGAAATTTCAAAACACA--AATTCAAATTAATCTATAGTGAATCAATTA 4315

Qy 365 TAGCCTTTCTGTAGGCAAGGTTAGTATCTTGTGTTTTTCC-----AGACCTCAAGGTGTTT 418
Db 4316 CAGGTTTTTGAATAACTTTTACTGTACTTGGTACTACTAATATTTAATTTAATA 4375

Qy 419 TTTGTTGTTTTTTTCATACCGGTGTGTGTGTGGTGTGGCCACTAAAAAGCTCAAGCA 478
Db 4376 ATTACATTAATAATAAATAATTAATTAATTTAAAAAATCTGCTTATATACCGATGC 4435

Qy 479 AGAATAATAACACTACACATACTAATAACCAATAGTATATAAAAAATAATAGCATCTGGC 538
Db 4436 ATTACTCTTTTCTAAGTTCAATATAGTAACAGTAACATAGGCGAAACCCCGCTCTTAC 4495

Qy 539 TAATTGCTGGACACTGTTTTAAGTGGTTTGCATGCCTCAGCTCATTAATCACTTACCTG 598
Db 4496 TAAAGTACAAAAATTAGCTGGGCGTGGTAGCGGTGCTGTAAACCTAGCTACTGGGAG 4555

Qy 599 TTATTATTGGCCCTATTTTACAAACAGGAGCCAGGCTCAGAGCAGTTAACTAACAGCC 658
Db 4556 GCAGAAACAGGAATCGCTTGAACCTGGAGGCGAGAGTTGCGAGTGAGCCAGATCGGC 4615

Qy 659 TCTCAAAAGAAACTCTCCAGAGATATTAATTTAAAAAATAATGAGAAATTAACACAC 718
Db 4616 CAGTGCAAGACTCCGCTTTCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGTAAAGAA 4675

Qy 719 AAGAAAGTTGAAATTTAGAGGTACAGGAGCTAAGCTTGTGTTGTTTGAACAGTGTCTG 778
Db 4676 AATATTTTATGAA-----AGATGCTCTGTGTTAATTTGTTATGATCAATATTTTTC 4730

Qy 779 CTACTGGGAAAGAGCGCAAGTCTTGGCTTTCTTAATAATTTGATACAGGACTCTCTGTAATC 838
Db 4731 CAATTAAGTGGAGGTCATAAGTATTTCC---AACATTGATCTAAATCCAGTGTAAACA 4787

Qy 839 ATATTTTGCATGCATGTAAGTAAGAAATGAAGCGGGTGCATGGCATCCAGTATCCAGTATC 898
Db 4788 AAAAGTTGATCTTTTATTAATAAATAAATTCAGGCGCGGCGCTCTGGCTCACGCTGCAATC 4847

Query Match	27.7%;	Score	277.4;	DB	4;	Length	115814;
Best Local Similarity	90.5%;	Pred. No.	2.6e-63;				
Matches	296;	Conservative	0;	Mismatches	31;	Indels	0;
Gaps	0;						
Qy	3	CAGAAGAAAGCACATCTCGGCGGGCGTGGTGGCTTCACACCTGTAATCCCAACACATTGG	62				
Db	8625	CAAAAAAAAAAAGATTCGGCGGGCGGCTGGCTCAGCGCTGTAATCCAGACACTTTGG	8566				
Qy	63	GAGGCCGAGGCGGTGGATCACTAGGTCAGGAGATGGAGACCATCTCTGGCTGACACGGTG	122				
Db	8565	GAGGCCGAGGCGGCGGATCAAGAGTTCAGGATCGAGACCATCTCTGGCTAACACGGTG	8506				
Qy	123	AAACCTGTCTCTACTAAAAACAAAAAATTAGCCGGGCGTGGTGCGAGGCGGCTGTAG	182				
Db	8505	AAACCCGCTCTCTACTAAAAATACAAAAAATTAGCCGGGCGTGGTGCGGCGGCTGTAG	8446				
Qy	183	TCCAGCTACTCTCGGAGGCTGAGCGAGGAGAAATGGCCTGAACCCGGGAGCGGACCTTGC	242				
Db	8445	TCCAGCTACTCTCGGAGGCTGAGCGAGGAGAAATGGCCTGAACCCGGGAGCGGAGCTTGC	8386				
Qy	243	AGTGAGCGGAGGTTGCAACCACTGCACCTCCAGCCTGGGCAACACAGTGAGAGCTCCGCTCA	302				
Db	8385	AGTGAGCGGAGATCGCGCCACTGCACCTCCAGCCTGGGCGGACAGAGCGAGACTCCGCTCA	8326				
Qy	303	AAAAAAAAAAAAAGAAAGAAAGAAAT	329				
Db	8325	AAAAAAAAAAAAAGAAAGAAAGAAAT	8299				

```

RESULT 15
US-09-949-016-190842/c
; Sequence 190842, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 190842
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-190842

```

Query Match	27.7%;	Score 276.8;	DB 4;	Length 601;
Best Local Similarity	89.0%;	Prod. No. 3.3e-64;		
Matches 299;	Conservative 0;	Mismatches 37;	Indels 0;	Gaps 0;
Qy	1	AACAGAAGAAAGC	CATCTCGGCGGGCGTGGTGCTCACACCTGTGTATTC	CAACACTTT 60
Db	471	AATTGAAGATATACAA	CTTGGCGGGCGGTGGCTCACGCTGTATTC	CAGCACTTT 412
Qy	61	GGGAGGCCGAGCGGGT	CGATCACTAGTCTCAGGAGATGGAGACCATCTCGCTC	GACACGG 120
Db	411	GGGAGGCCGAGCGGG	CGGATCACGAGTTCAGGAGATCGAGACCATCTCGCT	TACACAGG 352
Qy	121	TGAACCCCTGTCTCT	ACTATAAAAAACAAAAAATTAGCCGGCGGTGGTAGCGGCCCTGT	180
Db	351	TGAACCCCGCTCTCT	ACTATAAAAAACAAAAAATTAGCCGGCGGTGGTAGCGGCCCTGT	292
Qy	181	AGTCCCAGCTACT	CGGGAGGCTTAGGCGAGGAAATGGGCTTGAACCCGGGAGCGGAACTT	240

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 400.727 Seconds
(without alignments)
15667.668 Million cell updates/sec

Title: US-09-936-271c-13_copy_8000_9000
Perfect score: 1001
Sequence: 1 aacagagaagacatctc.....taaaataactatgttttt 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283.4	28.3	87467	19	US-10-741-601-5634 Sequence 5634, Ap
2	283.4	28.3	87467	21	US-10-741-600-17624 Sequence 17624, A
3	279.4	27.9	96599	14	US-10-105-948-4 Sequence 4, Appli
4	279.4	27.9	96599	18	US-10-052-482-178 Sequence 178, App
5	278.2	27.8	3407	10	US-09-764-872-749 Sequence 749, App
6	276.8	27.7	114793	15	US-10-148-806-3 Sequence 3, Appli
7	276.8	27.7	114793	22	US-10-859-792-3 Sequence 3, Appli

c 8	276.6	27.6	119501	17	US-10-174-319-15 Sequence 15, Appli
c 9	276.4	27.6	31474	10	US-09-764-891-8149 Sequence 8149, Ap
c 10	276.4	27.6	204621	13	US-10-087-192-958 Sequence 958, App
c 11	276.4	27.6	290367	20	US-10-719-993-6887 Sequence 6887, Ap
c 12	276.2	27.6	100445	19	US-10-322-281-170 Sequence 170, App
c 13	276	27.6	344805	20	US-10-779-271-1 Sequence 1, Appli
c 14	276	27.6	354592	22	US-10-737-082-70 Sequence 70, Appli
c 15	276	27.6	354592	22	US-10-765-790-70 Sequence 70, Appli
c 16	275.6	27.5	83361	19	US-10-741-601-5702 Sequence 5702, Ap
c 17	275.6	27.5	83361	21	US-10-741-600-17803 Sequence 17803, A
c 18	275.4	27.5	103730	21	US-10-741-600-17809 Sequence 17809, A
c 19	275.2	27.5	41104	9	US-09-816-685-3 Sequence 3, Appli
c 20	275.2	27.5	41104	18	US-10-639-708-3 Sequence 3, Appli
c 21	274.8	27.5	2328	10	US-09-764-891-6578 Sequence 6578, Ap
c 22	274.6	27.4	1400	9	US-09-263-959-295 Sequence 295, App
c 23	274.6	27.4	2325	10	US-09-764-891-6579 Sequence 6579, Ap
c 24	274.6	27.4	177380	21	US-10-484-577-683 Sequence 683, App
c 25	274.6	27.4	267156	11	US-09-968-007A-427 Sequence 427, App
c 26	274.6	27.4	267156	21	US-10-843-641A-6897 Sequence 6897, Ap
c 27	274.6	27.4	684973	9	US-09-263-959-1 Sequence 1, Appli
c 28	274.4	27.4	466	20	US-10-357-930-35244 Sequence 35244, A
c 29	274.2	27.4	14543	19	US-10-344-908-1 Sequence 1, Appli
c 30	274.2	27.4	14543	21	US-10-901-583-17 Sequence 17, Appli
c 31	273.6	27.3	392	20	US-10-357-930-40676 Sequence 40676, A
c 32	273.6	27.3	392	20	US-10-357-930-44084 Sequence 44084, A
c 33	273.6	27.3	466	13	US-10-027-632-114651 Sequence 114651, Sequence 114651,
c 34	273.6	27.3	466	17	US-10-027-632-114651 Sequence 76, Appli
c 35	273.6	27.3	43411	19	US-10-450-826-76 Sequence 66, Appli
c 36	273.4	27.3	67883	22	US-10-737-082-66 Sequence 66, Appli
c 37	273.4	27.3	67883	22	US-10-765-790-66 Sequence 66, Appli
c 38	273.4	27.3	202814	20	US-10-719-993-6812 Sequence 6812, Ap
c 39	273.2	27.3	93443	20	US-10-723-860-2089 Sequence 2089, Ap
c 40	273.2	27.3	276820	18	US-10-271-416-9 Sequence 9, Appli
c 41	273.2	27.3	338702	13	US-10-087-192-292 Sequence 292, App
c 42	273	27.3	303172	20	US-10-719-993-6890 Sequence 6890, Ap
c 43	273	27.3	322885	21	US-10-679-366-3 Sequence 3, Appli
c 44	272.8	27.3	32203	9	US-09-764-869-1849 Sequence 1849, Ap
c 45	272.8	27.3	32203	14	US-10-091-504-1849 Sequence 1849, Ap

ALIGNMENTS

RESULT 1
US-10-741-601-5634
; Sequence 5634, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5634
; LENGTH: 87467
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-5634

Query Match	28.3%	Score 283.4;	DB 19;	Length 87467;
Best Local Similarity	90.7%	Pred. No. 9.6e-61;		
Matches 302;	Conservative 0;	Mismatches 31;	Indels 0;	Gaps 0;
QY	1	AACAGAGAAGACATCTCGCCGGCGTGTGCTCACCTCTAATCCACACTTT	60	
Db	15709	AAAGTTGAAAGAAAAGATCGCCGGCGGTGCTCAACCTCTAATCCAGCACTTT	15768	
QY	61	GGGAGCCGAGCGGTGATCAGTCTAGGATCGAGACCATCTGCTGACACGG	120	
Db	15769	GGGAGCCGAGCGGTGATCAGTCTAGGATCGAGACCATCTGCTGACACGG	15828	

```
Qy 121 TGAACCCCTGCTCTACTATAAAACACAAAAATTAGCGGGCGTGTGGCAGGCGCCTGT 180
Db 15829 TGAACCCCGCTCTCTACTATAAAATACAAAAATTAGCGGGCGTGTGGCAGGCGCCTGT 15888
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGAGATGGCTGAACCCGGAGGCGGAACCTT 240
Db 15889 AGTCCAGCTACTCGGAGGCTGAGGCGAGAGATGGCTGAACCCGGAGGCGAGAGCTT 15948
Qy 241 GCAGTGAGCCGAGTTCGACCACTGCCTCAGCTGGCAACACAGTGAGACTCCGCTCT 300
Db 15949 GCAGTGAGCCGAGTTCGCGCCACTGCCTCAGCTGGGTACAGAGTGAGACTCCGCTCT 16008
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATACACA 333
Db 16009 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGA 16041

RESULT 2
US-10-741-600-17624
; Sequence 17624, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17624
; LENGTH: 87467
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17624

Query Match 28.3%; Score 283.4; DB 21; Length 87467;
Best Local Similarity 90.7%; Pred. No. 9.6e-61;
Matches 302; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 1 AACGAGAAAGCACATCTCGCGCGGCGTGGTGTGCTCACCTGTATATCCAAACACTTT 60
Db 15709 AAAAGTTGAAAGAAAAGATCGCGCGCGTGGTGTGCTCACCGCTGTATATCCAGCACTTT 15768
Qy 61 GGGAGCCGAGCGGCTGGATCCTAGTCTAGGAGATGGAGACCATCTGTGCTGCACACGG 120
Db 15769 GGGAGCCGAGCGGCTGGATCCTAGTCTAGGAGATGGAGACCATCTGTGCTGCACACGG 15828
Qy 121 TGAACCCCTGCTCTACTATAAAACACAAAAATTAGCGGGCGTGTGGCAGGCGCCTGT 180
Db 15829 TGAACCCCGCTCTCTACTATAAAATACAAAAATTAGCGGGCGTGTGGCAGGCGCCTGT 15888
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGAGATGGCTGAACCCGGAGGCGGAACCTT 240
Db 15889 AGTCCAGCTACTCGGAGGCTGAGGCGAGAGATGGCTGAACCCGGAGGCGAGAGCTT 15948
Qy 241 GCAGTGAGCCGAGTTCGACCACTGCCTCAGCTGGCAACACAGTGAGACTCCGCTCT 300
Db 15949 GCAGTGAGCCGAGTTCGCGCCACTGCCTCAGCTGGGTACAGAGTGAGACTCCGCTCT 16008
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATACACA 333
Db 16009 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGA 16041

RESULT 3
US-10-105-948-4
; Sequence 4, Application US/10105948
; Publication No. US20030064383A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
```

```
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; TITLE OF INVENTION: ASSOCIATED WITH ALTERED EXPRESSION OF PRLR
; FILE REFERENCE: 529452500127
; CURRENT APPLICATION NUMBER: US/10/105,948
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 96599
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-105-948-4

Query Match 27.9%; Score 279.4; DB 14; Length 96599;
Best Local Similarity 90.6%; Pred. No. 1e-59;
Matches 298; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 5 GAAGAAAGCACATCTCGCGCGGCGTGGTGTGCTCACACCTGTATATCCAAACACTTTGGGA 64
Db 49044 GAAATATACACATCTGGCGCGGCGGCTGCTCACGCTGTATATCCAGCACTTTGGGA 49103
Qy 65 GCGCGAGCGGCTGGATCCTAGTCTAGGAGATGGAGACCATCTGTGCTGCACCGTGAA 124
Db 49104 GCGCGAGCGGCGGATCAGAGTCTAGGAGATCGAGACCATCCCGGCTTAAACCGTGAA 49163
Qy 125 ACCCTGTCTCTACTATAAAACACAAAAATTAGCGGGCGTGTGGCAGGCGCCTGTAGTC 184
Db 49164 ACCCGTCTCTACTATAAAATACAAAAATTAGCGGGCGTGTGGCAGGCGCCTGTAGTC 49223
Qy 185 CCAGCTACTCGGAGGCTGAGGCGAGAGATGGCTGAAACCCGGAGGCGGAACCTTGCA 244
Db 49224 CCAGCTACTTGGAGGCTGAGGCGAGAGATGGCTGAAACCCGGAGGCGGAGCTTGCA 49283
Qy 245 TGAGCCGAGTTCGACCACTGCCTCCAGCTGGCAACACAGTGAGACTCCGCTCTCAA 304
Db 49284 TGAGCCGAGTTCGCGCCACTGCCTCCAGCTGGCGAGAGATGGCTGAAACCCGGAGGCGGAGCTTGCA 49343
Qy 305 AAAAAAAGAAAAAGAAAAAGAAAAATACACA 333
Db 49344 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 49372

RESULT 4
US-10-052-482-178
; Sequence 178, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 178
; LENGTH: 96599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-052-482-178

Query Match 27.9%; Score 279.4; DB 18; Length 96599;
Best Local Similarity 90.6%; Pred. No. 1e-59;
```

	Matches	298;	Conservative	0;	Mismatches	31;	Indels	0;	Gaps	0;
Qy	5	GA	AAAAAGCACATCTCGGCGGGCGTGGTGGCTCACACCTGTAATCCCAACACTTTGGGA	64						
Db	49044	GA	AAATAACACATCTGGGCGGGCGGGTGGCTCACGCGCTGTAATCCCAACACTTTGGGA	49103						
Qy	65	GG	CCGAGGGCGGTGGATCACTAGGTCAGAGAGTCGAGACCATCTTGGCTGACACCGTGAA	124						
Db	49104	GG	CCGAGGGCGGGGATACGAGGTCAGGAGATCGAGACCATCCCGGCTTAAACCGTGAA	49163						
Qy	125	AC	CTGCTCTACTAAAAACACAAAAAATAGCCGGCGTGTGCGACGGCGCTCTAGTC	184						
Db	49164	AC	CCCGTCTCTACTAAAAAATACAAAAAATAGCCGGCGCGTGTGCGCGCGCGCTCTAGTC	49223						
Qy	185	CA	GCTACTCGGAGGCTGAGGCAGAGAGTAATGGCTGAAACCCGGGAGCGCGAACTTGCAG	244						
Db	49224	CA	GCTACTTTGGAGGCTGAGGCAGAGAGTAATGGCTGAAACCCGGGAGCGCGAGCTTGCAG	49283						
Qy	245	TG	AGCCGAGGTTGCACCACTTGCACTCCAGCCCTGGGCAACACAGTGAGACTCCGCTCTCAA	304						
Db	49284	TG	AGCCGAGATTGCGCCACTTGCACTCCAGCCCTGGGCGACAGAGCGAGACTCCGCTCTCAA	49343						
Qy	305	AAAA	AAAAAAGAAAGAAAGAAATCACA	333						
Db	49344	AAAA	AAAAAAGAAAGAAAGAAAGAAATCACA	49372						

```

RESULT 5
US-09-764-872-749
; Sequence 749, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 749
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-749

```

```

RESULT 6
US-10-148-806-3
; Sequence 3, Application US/10148806
; Publication No. US20030138933A1
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148,806
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-148-806-3

```

```

RESULT 7
US-10-859-792-3
; Sequence 3, Application US/10859792
; Publication No. US20050136425A1
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; TITLE OF INVENTION: HELICASE
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/859,792
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US/10/148,806
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09

```

```
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-859-792-3

Query Match      27.7%; Score 276.8; DB 22; Length 114793;
Best Local Similarity 92.9%; Pred. No. 5.2e-59;
Matches 290; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 17 TCTCGCGCGGCGTGGCTCACACCTGTATATCCACACACTTTGGGAGCGGAGCGGG 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19284 TCTCGCGCGGCGGCTGGCTCACCGCTGTATATCCAGCACTTTGGGAGCGGAGCGGG 19343
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 77 TGGATCACTAGGTCAGGAGATGGAGACCATCTCTGGCTGACACGGTGAACCCCTGTCTTA 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19344 CGGATCAGAGTCAGGAGATCGAGACCATCTCTGGCTTAACACGGTGAACCCCTCTTA 19403
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 137 CTAACAAACAAAAAATTTAGCGGCGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGG 196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19404 CTAACAAATACAAAAATTTAGCGGCGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGG 19463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 197 GAGGCTGAGGAGGAGATGCGCTGACCCGAGGCGGAACTTGCAGTGAGCGAGGTT 256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19464 GAGGCTGAGGAGGAGATGCGCTGACCCGAGGCGGAGCTTGCAGTGAGCGGAGATC 19523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 257 GCACCACTGCACTCCAGCTCGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAA 316
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19524 GCGCACTGCACTCGAGCTGGGCGACAGCGAGACTCCGCTCTCAAAAAAAGAA 19583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 317 AAACAAAAAGAA 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19584 AAAAAAAGAA 19595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-174-319-15/c
; Sequence 15, Application US/10174319
; Publication No. US20030232771A1
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MARK3 EXPRESSION
; FILE REFERENCE: PTS-0018
; CURRENT APPLICATION NUMBER: US/10/174,319
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 121
; SEQ ID NO 15
; LENGTH: 119501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-174-319-15

Query Match      27.6%; Score 276.6; DB 17; Length 119501;
Best Local Similarity 91.0%; Pred. No. 6e-59;
Matches 294; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 2 ACAGAGAAAGACATCTCGCGCGGCGTGGTGGCTACACCTGTATATCCCAACACTTTG 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 62 GGAGCCGAGCGGCTGATCACTAGTTCAGGAGATGAGACCATCTCTGGCTGACCGGT 121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 67871 GGAGCCGAGCGGCGGATCACGAGGTTCAGGAGATGAGACCATCTCTGGCTTAAACCGGT 67812
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 122 GAAACCCCTGTCTTACTAAAAACAAAAAATTTAGCCGGCGTGGTGGCAGCGCGCTGTA 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-859-792-3

Query Match      27.7%; Score 276.8; DB 22; Length 114793;
Best Local Similarity 92.9%; Pred. No. 5.2e-59;
Matches 290; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 17 TCTCGCGCGGCGTGGCTCACACCTGTATATCCACACACTTTGGGAGCGGAGCGGG 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19284 TCTCGCGCGGCGGCTGGCTCACCGCTGTATATCCAGCACTTTGGGAGCGGAGCGGG 19343
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 77 TGGATCACTAGGTCAGGAGATGGAGACCATCTCTGGCTGACACGGTGAACCCCTGTCTTA 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19344 CGGATCAGAGTCAGGAGATCGAGACCATCTCTGGCTTAACACGGTGAACCCCTCTTA 19403
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 137 CTAACAAACAAAAAATTTAGCGGCGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGG 196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19404 CTAACAAATACAAAAATTTAGCGGCGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGG 19463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 197 GAGGCTGAGGAGGAGATGCGCTGACCCGAGGCGGAACTTGCAGTGAGCGAGGTT 256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19464 GAGGCTGAGGAGGAGATGCGCTGACCCGAGGCGGAGCTTGCAGTGAGCGGAGATC 19523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 257 GCACCACTGCACTCCAGCTCGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAA 316
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19524 GCGCACTGCACTCGAGCTGGGCGACAGCGAGACTCCGCTCTCAAAAAAAGAA 19583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 317 AAACAAAAAGAA 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19584 AAAAAAAGAA 19595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-09-764-891-8149/c
; Sequence 8149, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8149
; LENGTH: 31474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8149

Query Match      27.6%; Score 276.4; DB 10; Length 31474;
Best Local Similarity 89.2%; Pred. No. 3.7e-59;
Matches 298; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 6 AAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACCTGTAAATCCCAACACTTTGGGAG 65
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1304 AAAAAATATATTTTAGCGCGGCGGCTGAGCTACGCTGTAAATCCCAACACTTTGGGAG 1245
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 66 GCCGAGGCGGCTGATCACTAGGTGAGGATGAGACCATCTCTGGCTGACACCGTGAAA 125
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1244 GCCGAGGCGGCGATCACGAGGTGAGGATCGAGACCATCTCTGGCTTAAACCGTGAAA 1185
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 126 CCCTGTCTTACTAAAAACAAAAAATTAGCCGGCGTGGTGGCAGCGGCTGTAGTCC 185
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1184 CCCGCTCTTACTAAAAATACAAAAAATTAGCCGGCGTGGTGGCAGCGGCTGTAGTCC 1125
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 186 CAGTACTCGGAGGCTGAGGCGAGGAGATGGCTGACCCGGGAGCGGAACTTGCAGT 245
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1124 CAGTACTCGGAGGCTGAGGCGAGGAGATGGCTGACCCGGGAGCGGAGCTTGCAGT 1065
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 246 GAGCGGAGGTTGACCACTGCACTCCAGCTTGGGCAACACAGTGAGACTCCGCTCTCAAAA 305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1064 GAGCGGAGATCGCGCACTGCACTCCAGCTTGGGCGACAGAGCGAGACTCCGCTCTCAAAA 1005
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 306 AAAAAAAGAAAAAGAAAAATCATCTCAT 339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1004 AAAAAAAGAAAAAATTTTAT 971
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-10-087-192-958/c
; Sequence 958, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
```



```
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 958
; LENGTH: 204621
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(204621)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-958

Query Match 27.6%; Score 276.4; DB 13; Length 204621;
Best Local Similarity 90.5%; Pred. No. 8.5e-59;
Matches 295; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 2 ACAGAGAAAGCACATCTCGCGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 61
Db 23011 AAAGAACCTTGAACACTGCGCGGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 61
QY 62 GGAGCGCGAGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 121
Db 22951 GGAGCGCGAGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 121
QY 122 GAAACCCCTGCTCTACTTAAACACAAATAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 181
Db 22891 GAAACCCCTGCTCTACTTAAACACAAATAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 181
QY 182 GTCCAGCTACTCGGAGGCTGAGCAGGAGAAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 241
Db 22831 GTCCAGCTACTCGGAGGCTGAGCAGGAGAAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 241
QY 242 CAGTGAGCGGAGTTGACACCTGACCTCCAGCCCTGGGCAACACAGTGAGAGTCCGCTCTC 301
Db 22771 CAGTGAGCGGAGTTGACACCTGACCTCCAGCCCTGGGCAACACAGTGAGAGTCCGCTCTC 301
QY 302 AAAAAAAGAAAGAAAGAA 327
Db 22711 AAAAAAAGAAAGAAAGAA 22686

RESULT 11
US-10-719-993-6887/c
; Sequence 6887, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6887
; LENGTH: 290367
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(290367)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6887

Query Match 27.6%; Score 276.4; DB 20; Length 290367;
Best Local Similarity 90.5%; Pred. No. 1e-58;
Matches 295; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
```

```
QY 2 ACAGAGAAAGCACATCTCGCGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 61
Db 119843 AAAGAACCTTGAACACTGCGCGGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 119784
QY 62 GGAGCGCGAGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 121
Db 119783 GGAGCGCGAGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 119724
QY 122 GAAACCCCTGCTCTACTTAAACACAAATAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 181
Db 119723 GAAACCCCTGCTCTACTTAAACACAAATAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 119664
QY 182 GTCCAGCTACTCGGAGGCTGAGCAGGAGAAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 241
Db 119663 GTCCAGCTACTCGGAGGCTGAGCAGGAGAAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 119604
QY 242 CAGTGAGCGGAGTTGACACCTGACCTCCAGCCCTGGGCAACACAGTGAGAGTCCGCTCTC 301
Db 119603 CAGTGAGCGGAGTTGACACCTGACCTCCAGCCCTGGGCAACACAGTGAGAGTCCGCTCTC 119544
QY 302 AAAAAAAGAAAGAAAGAA 327
Db 119543 AAAAAAAGAAAGAAAGAA 119518

RESULT 12
US-10-322-281-170/c
; Sequence 170, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945201000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170
; LENGTH: 100445
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(100445)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-170

Query Match 27.6%; Score 276.2; DB 19; Length 100445;
Best Local Similarity 87.5%; Pred. No. 7e-59;
Matches 302; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
```

```
QY 6 AAGAAAGCACATCTCGCGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 65
Db 75610 AAGAAAGAACCTGAGCGGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 75551
QY 66 GCCGAGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 125
Db 75550 GCCGAGCGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 75491
QY 126 CCCTGTCTCTACTTAAACACAAATAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 185
Db 75490 CCCTGTCTCTACTTAAACACAAATAATAGCCGGGCTGATCTAGGTGAGAGATGAGAGACCATCTGGCTGACACCGT 75431
QY 186 CAGCTACTCGGAGGCTGAGCAGGAGAAATAGCCCTGAAACCCGGGAGGAGTTCAGT 245
Db 75430 CAGCTACTCGGAGGCTGAGCAGGAGAAATAGCCCTGAAACCCGGGAGGAGTTCAGT 75371
QY 246 GAGCGGAGTTGACACCTGACCTCCAGCCCTGGGCAACACAGTGAGAGTCCGCTCTCAAAA 305
Db 75370 GAGCGGAGTTGACACCTGACCTCCAGCCCTGGGCAACACAGTGAGAGTCCGCTCTCAAAA 75311
```

QY 306 AAAAAAAAAAGAAAAAGAAAAATCACATCTCATTTCAAGTGGTG 350
|||||
Db 75310 AAAAAAAAAAGAAAAAGAAAAATCCAGGCGAGTTGAAACAGTAG 75266
|||||

RESULT 13

US-10-779-271-1/c
; Sequence 1, Application US/10779271
; Publication No. US20040220387A1
; GENERAL INFORMATION:
; APPLICANT: AstraZeneca AB
; TITLE OF INVENTION: Methods
; FILE REFERENCE: AS2D-P02-251
; CURRENT APPLICATION NUMBER: US/10/779,271
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: 09/463,844
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: PCT/GB98/00259
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 9716162.4
; PRIOR FILING DATE: 1997-08-97
; PRIOR APPLICATION NUMBER: 60/535,986
; PRIOR FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 344805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-779-271-1

Query Match 27.6%; Score 276; DB 20; Length 344805;
Best Local Similarity 89.5%; Pred. No. 1.4e-58;
Matches 297; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 2 ACAGAGAAAGCACATCTCGCGCGCGGTGGTGTGCTGTAATTCACCAACTTTG 61
Db 88752 ACTTAAGAAATGGAATTCGCGCGCGGTGGTGTGCTGTAATTCACCAACTTTG 88693
|||||
QY 62 GGAGCCGAGCGGTGTGATCACTAGTTCAGGAGATGAGACCATCTCGGTGACACGGT 121
Db 88692 GGAGCCGAGCGCGGTGTGATCACTAGTTCAGGAGATGAGACCATCTCGGTGTAACCGT 88633
|||||
QY 122 GAAACCTGTCTCTACTAAAAACACAAAAATTAGCCGGCGTGTGCGAGCGCGCTGTA 181
Db 88632 GAAACCTGTCTCTACTAAAAATACAAAAATTAGCCGGCGTGTGCGAGCGCGCTGTA 88573
|||||
QY 182 GTCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTGAAACCCGGGAGCGGAACTTG 241
Db 88572 GTCCAGCTACTTGGAGGCTGAGCGAGGAGATGGCTGAAACCCGGGAGCGGAGCTTG 88513
|||||
QY 242 CAGTGAGCGAGGTGACACCTGCACTCCAGCTGGGCAACACAGTGAGACTCCGTCTC 301
Db 88512 CAGTGAGCGAGATTGCGCCACTGCACTCCAGCTGGGCGAGAGAGACTCCGTCTC 88453
|||||
QY 302 AAAAAAAAAAGAAAAAGAAATCACA 333
Db 88452 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 88421
|||||

RESULT 14

US-10-737-082-70/c
; Sequence 70, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

; FILE REFERENCE: 1657/2032
; CURRENT APPLICATION NUMBER: US/10/737,082
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 70
; LENGTH: 354592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-737-082-70

Query Match 27.6%; Score 276; DB 22; Length 354592;
Best Local Similarity 89.5%; Pred. No. 1.4e-58;
Matches 297; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 2 ACAGAGAAAGCACATCTCGCGCGCGGTGGTGTGCTGTAATTCACCAACTTTG 61
Db 91143 ACTTAAGAAATGGAATTCGCGCGCGGTGGTGTGCTGTAATTCACCAACTTTG 91084
|||||
QY 62 GGAGCCGAGCGGTGGATCACTAGTTCAGGAGATGAGACCATCTCGGTGACACGGT 121
Db 91083 GGAGCCGAGCGCGGTGGATCACTAGTTCAGGAGATGAGACCATCTCGGTGTAACCGT 91024
|||||
QY 122 GAAACCTGTCTCTACTAAAAACACAAAAATTAGCCGGCGTGTGCGAGCGCGCTGTA 181
Db 91023 GAAACCTGTCTCTACTAAAAATACAAAAATTAGCCGGCGTGTGCGAGCGCGCTGTA 90964
|||||
QY 182 GTCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTGAAACCCGGGAGCGGAACTTG 241
Db 90963 GTCCAGCTACTTGGGAGGCTGAGCGAGGAGATGGCTGAAACCCGGGAGCGGAGCTTG 90904
|||||
QY 242 CAGTGAGCCGAGGTGACACCTGCACTCCAGCTGGGCAACACAGTGAGACTCCGTCTC 301
Db 90903 CAGTGAGCCGAGATTGCGCCACTGCACTCCAGCTGGGCGAGAGAGACTCCGTCTC 90844
|||||
QY 302 AAAAAAAAAAGAAAAAGAAATCACA 333
Db 90843 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 90812
|||||

RESULT 15

US-10-765-790-70/c
; Sequence 70, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 70
; LENGTH: 354592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-70

Query Match 27.6%; Score 276; DB 22; Length 354592;
Best Local Similarity 89.5%; Pred. No. 1.4e-58;
Matches 297; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 2 ACAGAGAAAGCACATCTCGCGCGCGGTGGTGTGCTGTAATTCACCAACTTTG 61
|||||

Db 91143 ACTTAAGAAATGGAATTGCGGCGGGCGGGTGGGTCAAGCCCTGTAAATCCCGAGCACTTGG 91084
QY 62 GGAGGCCGAGCGGGTGGATCACTAGGTCAAGAGATGGAGACCATCCTGGCTTGACACGGT 121
Db 91083 GGAGGCCGAGCGGGCGGGATCAGAGGTCAAGAGATCGAGACCATCCCGGCTAAAAACGGT 91024
QY 122 GAAACCCCTGCTCTACTAAAAACACAAAAAATTAGCCGGCGGTGGTCAGGCGGCTGTA 181
Db 91023 GAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGCGGTAGTGGCGGGCGGCTGTA 90964
QY 182 GTCCAGCTACTCGGAGGCTGAGCAGGAGATGGCCCTGAACCCGGGAGGCGGAACCTTG 241
Db 90963 GTCCAGCTACTTGGAGGCTGAGCAGGAGATGGCGTGAACCCGGGAGGCGGAGCTTG 90904
QY 242 CAGTGAGCGGAGTTGCACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGTCTTC 301
Db 90903 CAGTGAGCGGAGATTGCGCCTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTTC 90844
QY 302 AAAAAAAAAAAGAAAGAAAGAAATCACA 333
Db 90843 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 90812

Search completed: July 4, 2005, 02:13:07
Job time : 406.727 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2005, 00:19:06 ; Search time 2176.56 Seconds

(without alignments)
17505.766 Million cell updates/sec

Title: US-09-936-271C-13_COPY_8000_9000

Perfect score: 1001

Sequence: 1 aacagaagaagcacatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	282	28.2	822	4	BG249643 602319736
2	277	27.7	759	9	AG185133 Pan trogl
3	274.2	27.4	564	5	BU860291 AGENCOURT
4	274.2	27.4	566	5	BU957747 AGENCOURT
5	273.6	27.3	691	4	BF965007 602268753
6	273.2	27.3	444	8	B89781 CIT-HSP-217
7	272.6	27.2	745	9	AG029211 Pan trogl
8	272	27.2	1073	8	BZ601227 WHACR89TR
9	271.8	27.2	664	8	AQ343449 RPC111-12
10	271.8	27.2	834	8	AQ744681 HS 5505 A
11	271	27.1	706	6	CA416075 UI-H-F50
12	270.4	27.0	709	9	AG167438 Pan trogl
13	270.2	27.0	517	1	AL046409 DKFp2434N
14	270	27.0	1051	8	BZ598731 WHACR86TR
15	269.8	27.0	731	8	AQ315280 RPC111-10
16	269.4	26.9	576	6	CA427630 UI-H-DFO
17	269	26.9	512	8	AQ395455 CITBI-EI-
18	268.8	26.9	375	2	AW303196 xtr89n02.x
19	268.8	26.9	513	5	BX502388 DKFp2779M
20	268.8	26.9	669	9	AG155696 Pan trogl
21	268.6	26.8	337	1	AI284640 qu23h09.x
22	268.6	26.8	469	6	CB217725 NISC nb04
23	268.2	26.8	513	5	BX442735 DKFp781P
24	268.2	26.8	541	6	CA437967 UI-H-DHO-

C	25	268.2	26.8	630	8	AQ628459 CITBI-EI-
C	26	268.2	26.8	715	5	BU617628 UI-H-DFO-
	27	268	26.8	662	9	AG144106 Pan trogl
	28	268	26.8	1050	6	CD252308 AGENCOURT
	29	267.8	26.8	379	2	AW301350 xs75g06.x
	30	267.8	26.8	514	8	AQ002001 CIT-HSP-2
	31	267.8	26.8	672	9	AG173650 Pan trogl
	32	267.6	26.7	515	8	AQ019249 CIT-HSP-2
	33	267.4	26.7	657	9	AG182991 Pan trogl
	34	267	26.7	585	4	BM479763 AGENCOURT
	35	267	26.7	647	4	BM559057 AGENCOURT
C	36	267	26.7	701	5	BU615970 UI-H-DFO-
	37	266.8	26.7	348	2	AW274349 xs74g09.x
C	38	266.8	26.7	400	5	BX482403 DKFp2686J
	39	266.8	26.7	581	2	BF677892 602084710
	40	266.4	26.6	612	7	CR817709 ib35d10.y
C	41	266.4	26.6	815	8	BZ608327 WHACR28TR
	42	266.2	26.6	666	9	AG075258 Pan trogl
C	43	266	26.6	650	9	AG116352 Pan trogl
	44	266	26.6	3680	3	HSM802780 Homo sapi
	45	265.8	26.6	354	5	BU963331 AGENCOURT

ALIGNMENTS

RESULT 1
BG249643
LOCUS 602319736P1 NIH_MGC_89 Homo sapiens CDNA clone IMAGE:4414905 5',
DEFINITION 822 bp mRNA linear EST 13-FEB-2001
rna sequence.
ACCESSION BG249643
VERSION BG249643.1 GI:12759459
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10143 row: h column: 10
High quality sequence stop: 546.
Location/Qualifiers
1..822
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4414905"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_89"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

FEATURES

Query Match 28.2%; Score 282; DB 4; Length 822;
Best Local Similarity 87.3%; Pred No. 7.4e-32;
Matches 309; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
ORIGIN

```
Qy 6 AAGAAAGCACATCTCGCGCGCGGTGGGTCTCAACCTGTAAATCCCAACACTTTGGGAG 65
Db 192 AAAATTCAGATATCGGCGCGCGGTGGGTCTACGCGCTGTAAATCCCAACACTTTGGGAG 251
Qy 66 GCCGAGGCGGTGGATCACTAGGTACAGAGATGGAGACCATCTCTGGCTGACACCGTGAAA 125
Db 252 GCCGAGGCGCGCGGTGGATCACTAGGTACAGAGATGGAGACCATCTCTGGCTGACACCGTGAAA 311
Qy 126 CCCTGTCTCTACTTAAACACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCC 185
Db 312 CCCCGTCTCTACTTAAACACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCC 371
Qy 186 CAGTACTCGGGAGGCTGAGCAGAGAAATGGCTTGAACCCGGAGGCGGAACTTGCAGT 245
Db 372 CAGTACTCGGGAGGCTGAGCAGAGAAATGGCTTGAACCCGGAGGCGGAGCTTGCAGT 431
Qy 246 GAGCGAGGTTGCACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGTCTCAAAA 305
Db 432 GAGCGAGATGGCGCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAA 491
Qy 306 AAAAAAAGAAAAAGAAAAAATCACATCTCATTTCAAGTGTGGCATTTAAAA 359
Db 492 AAAAAAAGAAAAAACAACAAAAAATAAAAAAAGGGGGGGGGGCAAAAA 545
```

```
RESULT 2
AG185133 759 bp DNA linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RP43-059C02.T7, genomic survey
DEFINITION
```

```
ACCESSION AG185133
VERSION AG185133.1 GI:16714813
```

```
KEYWORDS GSS
SOURCE Pan troglodytes (chimpanzee)
```

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library RPCI-43

TITLE

Unpublished

JOURNAL

2 (bases 1 to 759)

AUTHORS

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.

TITLE

Direct Submission

JOURNAL

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43. This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

Location/Qualifiers

FEATURES

source

1. 759

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="RP43-059C02.T7"

/sex="male"

/cell_type="lymphocytes"

/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 27.7%; Score 277; DB 9; Length 759;

Best Local Similarity 89.5%; Pred. No. 4.1e-31;

```
Matches 298; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 14 ACATCTCGCGCGGTGGGTCTCACACCTGTAAATCCCAACACTTTGGGAGGCGGAGGC 73
Db 286 AAAAATCGCGCGCGGTGGGTCTCACCTGTAAATCCCAACACTTTGGGAGGCGGAGGC 345
Qy 74 GGTGGATCACTAGGTGAGAGATGGAGACCATCTCTGGCTGACACCGTGAAACCTGTCT 133
Db 346 GGGCGGATCACGAGGTGAGAGATCGAGACCATCTCTGGCTTAAACACCGTGAAACCTGTCT 405
Qy 134 CTACTAAAAACACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCCCACTACT 193
Db 406 CTACTAAAAATACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCCCACTACT 465
Qy 194 CGGAGGCTGAGGAGAGAAATGGCTTGAACCCGGAGGCGGAACTTGCAGTGAGCCGAG 253
Db 466 CGGAGGCTGAGGAGAGAAATGGCTTGAACCCGGAGGCGGAGCTTGCAGTGAGCCGAG 525
Qy 254 GTTGCCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGTCTCAAAAAAATAAAA 313
Db 526 GTTGTCCTTGCCTCCAGCCTGGGCGACAGAGTGAGACTCCGTCTCAAAAAAATAAAA 585
Qy 314 AGAAAGAAAGAAATACATCTCTCATTTCAAGTG 346
Db 586 AAAAAATCAATAAGTGAATGACTTGGGTG 618
```

RESULT 3

BU860291/c

LOCUS

DEFINITION

AGENCOURT 10440048 NIH MGC 107 Homo sapiens cDNA clone

IMAGE: 6651369 5', mRNA sequence.

ACCESSION BU860291

VERSION BU860291.1 GI:24045283

KEYWORDS EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: Agencourt Bioscience Corporation

Cloned through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM2898 row: b column: 09

High quality sequence stop: 543.

FEATURES

source

1. 564

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6651369"

/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_107"

/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;

Site 2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match	27.4%;	Score 274.2;	DB 5;	Length 564;
Best Local Similarity	91.2%;	Prod. No. 1.2e-30;		
Matches	291;	Conservative	0;	Mismatches 28; Indels 0; Gaps 0;
17	TCTCGCGCGGGCGTGGCTCACACCTGTAATCCCAACACTTTTGGAGCGCCGAGCGGG	76		
328	TTTTCGGCCGGCGCGTGGCTCACGCCCTGTAATCCAGCACTTTTGGAGCGCCGAGCGGG	269		
77	TGGATCACTAGGTCAAGGAGATGGAGACCATCTGTGCTGCACACGGTGAAACCCCTGTCTCTA	136		
268	TGGATCACAGGTCAAGGAGATCGAGACCATCTGTGCTTAACACGGTGAAACCCCGTCTCTA	209		
137	CTAATAACACAAAAAATTAGCGGGCGTGGTGACAGGCGCCTGTAGTCCGAGCTACTCGG	196		
208	CTAATAATACAAAAAATTAGCGGGCGTGGTGACAGGCGCCTGTAGTCCGAGCTACTCGG	149		
197	GAGGCTGAGGCGAGGAGATGGCTCAACCCGGAGGCGGAACCTTGCAGTGTAGCCGAGGTT	256		
148	GAGGCTGAGGCGAGGAGATGGCGTGAACCCGGAGGCGGAGCTTGCAGTGTAGCCGAGATC	89		
257	GCACCACCTGCATCTCCAGCGCTGGGCAACACAGTGTAGACTCCGCTCTCAAAAAA	316		
88	GGCCACTGCATCTCAGCGCTGGGCGACAGAGCGAGACTCCGCTCTCAAAAAA	29		
317	AAAGAAAAGAAATCACATC	335		
28	AAAAACTCGAGACTAGTTTC	10		

```

RESULT 4
BU957747/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU957747 566 bp mRNA linear EST 21-OCT-2002
AGENCORE10620302 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6731834 5', mRNA sequence.
BU957747
BU957747.1 GI:24187319
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Makmyota; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 566)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM3058 row: o column: 01
High quality sequence stop: 450.
Location/Qualifiers
1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6731834"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTS7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

ORIGIN	Query Match	Best Local Similarity	27.4%; Score 274.2; DB 5; Length 566;
	Matches	291; Conservative	0; Mismatches 28; Indels 0; Gaps 0;
Qy	17	TCTCGCGCGCGTGGTGGCTCACACTGTAATCCCAACACTTTGGAGGCCGAGCGGG	76
Db	328	TTTTCGCGCGCGCGTGGTGGCTCACGCGTGTAAATCCGAGCACATTTGGAGGCCGAGCGGG	269
Qy	77	TGGATCACTAGGTTCAGGAGATGAGAGACCATCTCTGGCTGACACCGGTGAAACCCCTGTCTCTA	136
Db	268	TGATCAGAGGTTCAGGAGATCGAGACCATCTCTGGCTTAACACGCTGAACCCCGTCTCTA	209
Qy	137	CTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGCGCGCTGTAGTCCGAGCTACTCGG	196
Db	208	CTAAAAATACAAAAAATTAGCCGGCGGTGGTGGCGGGCGCTGTAGTCCGAGCTACTCGG	149
Qy	197	GAGCTGAGCAGAGAAATGGCTGTAAACCGGAGGCGGAACTTGCAGTGAGCGGAGTT	256
Db	148	GAGGCTGAGCAGAGAAATGGCGTGAACCGGAGGCGGAGCTTGCAGTGAGCGGAGATC	89
Qy	257	GCACCACTGCATCCAGCCTGGGCAACACAGTCGAGACTCGCTCTCAAAAAAAGAA	316
Db	88	CGCGCACTGCATCCAGCCTGGGCGACAGCGAGACTCGCTCTCAAAAAAAGAA	29
Qy	317	AAAGAAAGAAATCACATC 335	
Db	28	AAAACTCGAGACTAGTTC 10	
RESULT 5	BF965007	602268753F1 NIH_MGC_81 Homo sapiens	691 bp mRNA linear EST 22-JAN-2001
LOCUS	BF965007	602268753F1 NIH_MGC_81 Homo sapiens	691 bp mRNA linear EST 22-JAN-2001
DEFINITION	BF965007	602268753F1 NIH_MGC_81 Homo sapiens	691 bp mRNA linear EST 22-JAN-2001
ACCESSION	BF965007	602268753F1 NIH_MGC_81 Homo sapiens	691 bp mRNA linear EST 22-JAN-2001
VERSION	BF965007.1	GI:12332222	691 bp mRNA linear EST 22-JAN-2001
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1. (bases 1 to 691)		
JOURNAL	NTM-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: rgs@nih.gov		
	Tissue Procurement: CLONTECH Laboratories, Inc.		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: L12M1217 row: j column: 14		
	High quality sequence stop: 445.		

FEATURES
SOURCE

```

High quality sequence stop: 445.
Location/Qualifiers
1. .691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4356973"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 81"
/notes="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI
(ggccattagcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCCATATGCCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCCGACATG-dT(30)BN-3' (where B = A,
C, G and N = A, C, G, or T). Average insert size

```

1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```
ORIGIN
Query Match      27.3%; Score 273.6; DB 4; Length 691;
Best Local Similarity 89.6%; Pred. No. 1.3e-30;
Matches 294; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1 AACAGNAGAACACATCTCGCGCGGCGTGGTGCCTCACACCTGTATATCCACACTTTT 60
Db 132 AAAAGTTGAAGAAAAGATAGCGCGGCGGCTGCTCACCGCTGTATATCCAGACTTT 191
Qy 61 GGGAGGCGCGGCGGCGTGGATCACTAGGTCAAGGATGAGAGACCATCTGCTGCACACGG 120
Db 192 GGGAGGCCAAGCGGGTGTATCAAGGTCAAGGATGAGAGACCATCTGCTGCACACGG 251
Qy 121 TGAACCCCTGTCTTACTATAAAACACAAAAATTAGCGCGGCGTGGTGGCAGGCGCCTGT 180
Db 252 TGAACCCCGTCTTACTATAAAATACAAAAATTAGCGCGGCGTGGTGGCAGGCTGT 311
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGACGAGGATGGCTGAACCGGAGCGCGACTT 240
Db 312 AGTCCAGCTACTCGGAGGCTGAGGACGAGGATGGCGTGAACCGGAGCGCAGAGCTT 371
Qy 241 GCAGTGAGCGGAGGTTGCACCACTGCACCTCCAGCCTGGGCAACACAGTGAGACTCCGCT 300
Db 372 GCAGTGAGCGGAGGTTGGCCACTGCACTCCAGCTGGGTGACAGTGAGACTCCGCTCC 431
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAA 328
Db 432 CAAAAAAGAAAAAGAAAAAGAAAA 459

RESULT 6
B89781
LOCUS      B89781
DEFINITION CIT-HSP-2171J2.TF CIT-HSP Homo sapiens genomic clone 2171J2,
genomic survey sequence.
ACCESSION  B89781
VERSION    B89781.1 GI:2972261
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Other GSSs: CIT-HSP-2171J2.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamads@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..444
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7103508"
                     /db_xref="taxon:9606"

ORIGIN
Query Match      27.3%; Score 273.2; DB 8; Length 444;
Best Local Similarity 90.6%; Pred. No. 1.8e-30;
Matches 290; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 9 AAAGCACATCTCGCGCGGCGTGGTGCCTCACACTGTATATCCAACTTTTGGAGGCC 68
Db 60 AAAGCTCAATAGCGCGGCGGCTGCTCAGCGTGTATATCCAGCACTTTGGAGGCC 119
Qy 69 GAGCGGGTGGATCACTAGGTCAAGGATGGAGACCATCTGCTGCACACGTTGAAACCC 128
Db 120 GAGCGGGTGGATCACAGGTCAAGGATCGAGACCATCTGCTGTAAACCGTGAACCCC 179
Qy 129 TGTCTCTACTAAAAACACAAAAATTAGCGCGGCGTGGTGGCAGCGCTGTAGTCCCAG 188
Db 180 CGTCTCTACTAAAAATACAAAAATTAGCGCGGCGTGGTGGCAGCGCTGTAGTCCCAG 239
Qy 189 CTACTCGGGAGGTGAGCGCAGGAGAAATGGCTGAACCCGGGAGCGGAACTTGCAGTGAG 248
Db 240 CTACTCGGGAGGTGAGCGCAGGAGAAATGGCTGAACCCGGGAGCGGAACTTGCAGTGAG 299
Qy 249 CCGAGGTTGACCACTGCACCTCCAGCTGGGCAACACAGTGAGACTCCGTTCAAAAAA 308
Db 300 CCGAGATCACGCCACTGCACCTCCAGCTGGGCGCAGCAGCGAGACTCCGTTCAAAAAA 359
Qy 309 AAAAAAGAAAAAGAAAAAGAAA 328
Db 360 AAAAAAGAAAAAGAAAAAGAAA 379

RESULT 7
AG029211/c
LOCUS      AG029211
DEFINITION Pan troglodytes DNA, clone: PTB-001D16.F, genomic survey sequence.
ACCESSION  AG029211
VERSION    AG029211.1 GI:16556083
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 745)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-cho,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbres@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1..745

FEATURES             Location/Qualifiers
     source           1..745
```



```
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-001D16.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
```

ORIGIN

```
Query Match      27.2%; Score 272.6; DB 9; Length 745;
Best Local Similarity 89.6%; Pred. No. 1.8e-30;
Matches 293; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2 ACAGAGAAAGACATCTCGCGCGCGTGGTGCACACCTGTAATCCCAACATTG 61
   |||||
Db 415 ATAGAAAGAGCTCCTCGCGCGCGTGGTGCACCTGTAATCCCAACATTG 356
   |||||

QY 62 GGAGCCGAGCGGTGGATCACTAGTTCAGGAGATGAGACCATCTCTGGCTGACACGGT 121
   |||||
Db 355 GGAGCCGAGGTGGCGGATCAGAGGTTCAGGAGATCGAGACCATCTCTGGCTGACACGGT 296
   |||||

QY 122 GAAACCTGTCTCTACTAAAAACAAAAATTAGCCGGCGTGGTGCAGCGCGCTGTA 181
   |||||
Db 295 GAAACCTGTCTCTACTAAAAACAAAAATTAGCCGGCGTGGTGCAGCGCGCTGTA 236
   |||||

QY 182 GTCCAGCTACTCGGAGGCTGAGCAGGAGATGGCTGAAACCGGAGCGGAACTTG 241
   |||||
Db 235 GTCCAGCTACTCGGAGGCTGAGCAGGAGATGGCTGAAACCGGAGCGGAACTTG 176
   |||||

QY 242 CAGTGAGCGGAGGTGACACCTGACCTCCAGCTGGGCAACACAGTCAGACTCCGCTC 301
   |||||
Db 175 CAGTGAGCGGAGATCGGCCACTGCACTCCAGCTGGGCGAGAGAGCTCGCTC 116
   |||||

QY 302 AAAAAAAGAAAAAGAAAAAGAAA 328
   |||||
Db 115 AAAAAAAGAAAAAGAAAAAGAAA 89
   |||||
```

RESULT 8

```
BZ601227
LOCUS WHADC9TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
DEFINITION sapiens genomic clone MCF7_1-19p9, genomic survey sequence.
ACCESSION BZ601227
VERSION BZ601227.1 GI:31509689
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1073)
Volik, S., Zhao, S., Chin, K., Brebner, J. H., Herndon, D. R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J. W., and Collins, C.

End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)

JOURNAL

MEDLINE

PUBMED

12788976

Contact: Volik SV

Colin Collins' lab

UCSF Comprehensive Cancer Center

UCSF Box 0808, San Francisco, CA 94143-0808, USA

Tel: 415 502 7066

Fax: 415 502 5665

Email: svolik@cc.ucsf.edu

This clone is available from Amplicon Express

http://www.genomex.com

Class: BAC ends.

FEATURES

source

1..1073

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

```
/clone="MCF7_1-19p9"
```

```
/sex="female"
```

```
/clone_lib="Human MCF7 breast cancer cell line library
```

```
(MCF7.1)"
```

/note="vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

```
Query Match      27.2%; Score 272; DB 8; Length 1073;
Best Local Similarity 89.3%; Pred. No. 1.9e-30;
Matches 293; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 6 AAGAAACACATCTCGCGCGCGTGGTGCACACCTGTAATCCCAACATTG 65
   |||||
Db 318 AAAAAAAGAAAAAGAAAAAGAAA 377
   |||||

QY 66 GCCGAGCGGTGATCACTAGTTCAGGAGATGAGACCATCTCTGGCTGACACGGTGA 125
   |||||
Db 378 GCCGAGCGGTGATCACTAGTTCAGGAGATGAGACCATCTCTGGCTGACACGGTGA 437
   |||||

QY 126 CCTGTCTCTACTAAAAACAAAAATTAGCCGGCGTGGTGCAGCGCTGTAGTCC 185
   |||||
Db 438 CCTGTCTCTACTAAAAACAAAAATTAGCCGGCGTGGTGCAGCGCTGTAGTCC 497
   |||||

QY 186 CAGTACTCGGAGGCTGAGCAGGAGATGGCTGAAACCGGAGCGGAACTTGCA 245
   |||||
Db 498 CAGTACTCGGAGGCTGAGCAGGAGATGGCTGAAACCGGAGCGGAACTTGCA 557
   |||||

QY 246 GAGCGAGGTGACACCTGCACTCCAGCTGGGCAACACAGTCAGACTCCGCTCA 305
   |||||
Db 558 GAGCGAGGTGACACCTGCACTCCAGCTGGGCGAGAGAGCTCCGCTCA 617
   |||||

QY 306 AAAAAAAGAAAAAGAAAAAGAAA 333
   |||||
Db 618 AAAAAAAGAAAAAGAAAAAGAAA 645
   |||||
```

RESULT 9

```
AQ343449
LOCUS RPI11-122K4.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-122K4,
DEFINITION genomic survey sequence.
ACCESSION AQ343449
VERSION AQ343449.1 GI:4168345
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 664)
Zhao, S., Adams, M. D., Nierman, W., Malek, J., de Jong, P., and Venter, J. C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL

COMMENT

Unpublished (1997)

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

FEATURES

```

source
1. .664
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7546707"
/db_xref="taxon:9606"
/clones="RPCI-11-122K4"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"

ORIGIN
Query Match 27.2%; Score 271.8; DB 8; Length 664;
Best Local Similarity 90.1%; Pred. No. 2.5e-30;
Matches 291; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 21 GCGCGGCGGTGGCTCACACCTGTATATCCACACATTTGGAGGCGGAGCGGTGGA 80
Db 42 GCGCGGCGGCTGGCTCACGCTGTATATCCAGCACTTTGGAGGCGGAGCGGCGGA 101
Qy 81 TCACGTAGTCAGGAGTGGAGACCATCTGGCTGACACGGTGAAACCTGTCTACTAA 140
Db 102 TCACGTAGTCAGGAGTGGAGACCATCTGGCTTAACACGGTGAAACCCGCTCTACTAA 161
Qy 141 AAACACAAAAAATTAGCGGCGGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGGAGG 200
Db 162 AAATACAAAAAATTAGCAGCGGTGGTAGCGGCGCTGTAGTCCAGCTACTCGGAGG 221
Qy 201 CTGAGGCGAGGAGATGGCTGAAACCCGGGCGCGAACTTCAGTGAGCGCGAGTTGCAC 260
Db 222 CTGAGGCGAGGAGATGGCTGAAACCCGGGCGCGAGCTTCAGTGAGCGCGAGATCGCG 281
Qy 261 CACTGCATCTCAGCTGGGCAACAGTGAAGTCCGCTCAAAAAAAGAAAGAAAG 320
Db 282 CACTGCATCTCAGCTGGGCAACAGTGAAGTCCGCTCAAAAAAAGAAAGAAAG 341
Qy 321 AAAAGAAATCATCTCAATCAA 343
Db 342 AAAAAGGCTGAGTTGTTCAA 364

RESULT 10
AQ744681/c
LOCUS
DEFINITION HS_5505_A2_B04_SP6_RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plates=1081 Col=8 Row=C, genomic survey sequence.
ACCESSION AQ744681
VERSION AQ744681.1 GI:5522203
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 834)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACpac Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1081 row: C column: 8
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 834.
Location/Qualifiers
source
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1081 Col=8 Row=C"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN
Query Match 27.2%; Score 271.8; DB 8; Length 834;
Best Local Similarity 88.8%; Pred. No. 2.3e-30;
Matches 294; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 9 AAAGCACATCTCGGCGGCGGTGGCTCACACCTGTATATCCACACATTTGGAGGCC 68
Db 544 AAATCAACATAGCGCGGCGGTGGCTCATGCTGTATATCCAGCACATTTGGAGGCT 485
Qy 69 GAGCGGCTGGATCCTAGGTGAGGAGATGAGACCATCTCTGGCTGACACGGTCAAAACCC 128
Db 484 GAGCGGCGGATCAGAGGTGAGGAGATGAGACCATCTCTGGCTGACACGGTCAAGGCC 425
Qy 129 TGTCTTCTACTAAACACAAAAAATTAGCCGGCGGTGGTGGCAGCGCTGTAGTCCAG 188
Db 424 CGTCTCTACTAAAAAATTAGCCGGCGGAGGTGGTGGCGCTGTAGTCCAG 365
Qy 189 CTACTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGGCGGAGCTTGCAGTGCAG 248
Db 364 CTACTTGGGAGGCTGAGGAGGAGATGGCGTGAACCCGGGCGGAGCTTGCAGTGCAG 305
Qy 249 CCGAGGTTGCACCACTCCTCCAGCTGGGCAACACAGTGCAGACTCCGCTCAAAAAAA 308
Db 304 CCGAGATCGCGCCACTGCACTCCAGCTGGGCAACAGTGCAGACTCCGCTCAAAAAAA 245
Qy 309 AAAAAAGAAAGAAAGAAATCATCTCAT 339
Db 244 AAAAAAAGAAAGAAAGAAATCATCTCAT 214

RESULT 11
CA416075/c
LOCUS
DEFINITION UI-H-PEO-bbs-o-16-0-UI.s1 NCI CGAP PE0 Homo sapiens cDNA clone
UI-H-PEO-bbs-o-16-0-UI 3', mRNA sequence.
ACCESSION CA416075
VERSION CA416075.1 GI:24778726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

```


DKFP2p434N147 3', mRNA sequence.
AL046409
VERSION AL046409.1 GI:5434485
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 517)
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Koehler, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BPFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
r1 sequence also available.
This clone (DKFP2p434N147) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..517
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP2p434N147"
/tissue type="testis"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 27.0%; Score 270.2; DB 1; Length 517;
Best Local Similarity 89.8%; Pred. No. 4.6e-30; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 33;
Qy 6 AAGAAAGCACATCTCGGCGGGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 65
Db 325 AAAAAAGTATTAGGCGGGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 266
Qy 66 GCCGAGGCGGTGTATCTAGTGTGAGAGATGAGACCATCTCTGGTGTGACACGGTGAAA 125
Db 265 GCCGAGGCGGGATCATCAGAGGTGAGAGATGAGACCATCTCTGGTGTGAAAACGGTGAAA 206
Qy 126 CCCTGTCTTACTAAAAACAAAAAATTAGCCGGCGGTGGTGTGAGCGCGCTGTAGTCC 185
Db 205 CCGCGTCTTACTAAAAATACAAAAAATTAGCCGGCGGTGGTGTGAGCGCGCTGTAGTCC 146
Qy 186 CAGTCTACTCGGAGGCTGAGGAGAGATGGCTGAACCCGGGAGGCGGAACCTTGCACT 245
Db 145 CAGTCTACTCGGAGGCTGAGGAGAGATGGCTGAACCCGGGAGGCGGAACCTTGCACT 86
Qy 246 GAGCGGAGTTGACACCTGCACTCCAGCTGGGGCAACACAGTGAGACTCCGCTCTCAAAA 305
Db 85 GAGCGGAGATCCCGCACTGCACTCCAGCTGGGGCAACAGAGCAAGACTCCGCTCTCAAAC 26
Qy 306 AAAAAAAGAAAAAGAAAAA 328
Db 25 AAAAAAAGAAAAAGAAAAA 3

RESULT 14

BZ598731
LOCUS BZ598731 1051 bp DNA linear GSS 08-JUN-2003
DEFINITION WHACR86TR Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-1704, genomic survey sequence.
ACCESSION BZ598731

BZ598731.1 GI:31507193
GSS.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1051)
Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J.W. and Collins, C.
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
MEDLINE 22709111
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
Location/Qualifiers
1..1051
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-1704"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/note="Vector: pECBAC1; Site_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

FEATURES

source
1..1051
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-1704"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/note="Vector: pECBAC1; Site_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 27.0%; Score 270; DB 8; Length 1051;
Best Local Similarity 86.8%; Pred. No. 3.8e-30; Indels 0; Gaps 0;
Matches 297; Conservative 0; Mismatches 45;
Qy 6 AAGAAAGCACATCTCGGCGGGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 65
Db 226 AAAAAAGCACTGATCGGCTGGGCTGGTGTACACCTGTAAATCCCAACACTTTGGAG 285
Qy 66 GCCGAGGCGGTGGATCACTAGGTGAGAGATGAGACCATCTCTGGTGTGACACGGTGAAA 125
Db 286 GCTGAGGCGAGTGGATCAGGAGGTGAGAGAACGAGACCATCTCTGGTGTGAAAACGGTGAAA 345
Qy 126 CCCTGTCTTACTAAAAACAAAAAATTAGCCGGCGGTGGTGTGAGCGCGCTGTAGTCC 185
Db 346 CCGCGTCTTACTAAAAATACAAAAAATTAGTGGGCGGTGGTGTGAGCGACCTGTAGTCC 405
Qy 186 CAGTCTACTCGGAGGCTGAGGAGAGATGGCTGAACCCGGGAGGCGGAACCTTGCACT 245
Db 406 CAGTCTACTCGGAGGCTGAGGAGAGATGGCTGAGCCAGGAGGCGGAGCTTGCACT 465
Qy 246 GAGCGGAGTTGACACCTGCACTCCAGCTGGGGCAACACAGTGAGACTCCGCTCTCAAAA 305
Db 466 GAGCGGAGATCAGCCCACTGCACTCCAGCTGGGGCAACAGAGTGAGACTCCGCTCTCAAAA 525
Qy 306 AAAAAAAGAAAAAGAAAAATCAGATCTCATTCAAGTGG 347
Db 526 AAAAAAAGAAAAAGAAAAATCAGATCTTATGATTTG 567

RESULT 15

BZ598731
LOCUS BZ598731 731 bp DNA linear GSS 04-MAY-1999
DEFINITION RPC111-104N24.TV RPC1-11 Homo sapiens genomic clone RPC1-11-104N24, genomic survey sequence.

ACCESSION AQ315280
VERSION AQ315280.1 GI:4046743
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 731)
Adams M.D., Rounsley S.D., Zhao S., Bass S., Linher K., Golden K.,
Berry K., Granger D., Suh E., Wible C., de Jong P. and Venter J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPC111-104N24.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbs@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@jomb.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: 17
Class: BAC ends.
FEATURES
Location/Qualifiers
1..731
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:753987"
/db_xref="taxon:9606"
/clone="RPCI-11-104N24"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

ORIGIN
Query Match 27.0%; Score 269.8; DB 8; Length 731;
Best Local Similarity 90.0%; Pred. No. 4.7e-30;
Matches 289; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 6 AAGAAAGCACATCTCGGCGGCGGTGGTGGCTCACACCTGTAATCCCAACACTTTGGGAG 65
DB 124 AAGAAAGAGGTTTAGCGCGGCGGTGGTGGCTCACACCTGTAATCCCAACACTTTGGGAG 183
QY 66 GCCGAGCGGTGATCACTAGGTGAGGATGAGACCATCTGGCTGACACCGTGAAA 125
DB 184 GCCGAGCGGCGGATCACGAGGTGAGGATGAGACCATCTGGCTGAAAACCGTGAAA 243
QY 126 CCCTGTCTCTACTAAACACAAAATTTAGCGGCGGTGGTGGCGGCGCTGTAGTCC 185
DB 244 CCCCCTCTCTACTAAATAACAAAATTTAGCGGCGGTGGTGGCGGCGCTGTAGTCC 303
QY 186 CAGTACTCTCGGAGGCTCAGGAGGAGATGGCTGAACCCGGGAGGCGGAACCTTGCACT 245
DB 304 CAGTACTCTCGGAGGCTCAGGAGGAGATGGCTGAACCCGGGAGGCTTGCACT 363
QY 246 GAGCGAGGTTGACCACTGACCTCCAGCTGGGCAACACAGTGAAGTCCGTCTCAAAA 305
DB 364 GAGCGAGATCCCGCACTGCACTCCAGCTGGGCGACAGAGCGAGACTCCGTCTCAAAA 423
QY 306 AAAAAAAGAAAAGAAGA 326
DB 424 AAAAAAAGAAAAGAAGA 444

Search completed: July 4, 2005, 15:07:14
Job time : 2183.56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 15:07:24 ; Search time 73 Seconds
(without alignments)
1552.340 Million cell updates/sec

Title: US-09-936-271c-14

Perfect score: 1608

Sequence: 1 MATARPPPMWLALITALL.....VVTNLCKFTKWIQTIQANS 293

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_16Dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	293	2	Aay16777 Human ker
2	1608	100.0	293	2	Aay30524 Human PDS
3	1608	100.0	293	2	Aay38412 Human sec
4	1608	100.0	293	2	Aay38426 Human sec
5	1608	100.0	293	3	Aay66726 Membrane-
6	1608	100.0	293	3	Aab21296 Human KLIK
7	1608	100.0	293	4	Aau12399 Human PRO
8	1608	100.0	293	4	Aab65249 Human PRO
9	1608	100.0	293	5	Aau81966 Human PRO
10	1608	100.0	293	6	Abu58064 Human PRO
11	1608	100.0	293	6	Abu59142 Novel hum
12	1608	100.0	293	6	Abu82654 Human sec
13	1608	100.0	293	6	Abu17843 Novel hum
14	1608	100.0	293	6	Ada57425 Human sec
15	1608	100.0	293	6	Ada56974 Human sec
16	1608	100.0	293	6	Ada57427 Human sec
17	1608	100.0	293	6	Ada57428 Human sec
18	1608	100.0	293	6	Abu60573 Human sec
19	1608	100.0	293	6	Abu13955 Human PRO
20	1608	100.0	293	6	Abu81097 Human PRO
21	1608	100.0	293	6	Abu72540 Novel hum
22	1608	100.0	293	6	Abu66797 Human PRO
23	1608	100.0	293	6	Ada41303 Human sec
24	1608	100.0	293	6	Ada41305 Human sec
25	1608	100.0	293	6	Ada40825 Human sec

26	1608	100.0	293	6	ADA41306 Human sec
27	1608	100.0	293	6	ABU59878 Novel sec
28	1608	100.0	293	6	ABU59289 Human PRO
29	1608	100.0	293	6	ABO25986 Human PRO
30	1608	100.0	293	6	ABO25068 Human sec
31	1608	100.0	293	6	ABU58995 Human sec
32	1608	100.0	293	6	ABU92373 Novel hum
33	1608	100.0	293	6	ABU59438 Novel hum
34	1608	100.0	293	6	ABU67073 Human sec
35	1608	100.0	293	6	ABU92204 Novel hum
36	1608	100.0	293	6	ABU10910 Human PRO
37	1608	100.0	293	6	ABU81662 Novel hum
38	1608	100.0	293	6	ABU88601 Human PRO
39	1608	100.0	293	6	ABO34115 Human PRO
40	1608	100.0	293	6	ADA45975 Novel hum
41	1608	100.0	293	6	ADA76406 Human PRO
42	1608	100.0	293	6	ADA19056 Human PRO
43	1608	100.0	293	6	ADA61679 Homo sapi
44	1608	100.0	293	6	ADB19464 Novel hum
45	1608	100.0	293	6	ADB28005 Human PRO

ALIGNMENTS

RESULT 1

AY16777
ID AY16777 standard; protein; 293 AA.

XX AC AY16777;
XX DT 02-AUG-1999 (first entry)
XX DE Human keratinocyte derived protease (KDP).
XX KW Keratinocyte derived protease; KDP; protease; keratinocyte; human;
XX KW skin care product; skin flaking; dandruff; laundry detergent composition;
XX KW cleaning composition; dishwashing product.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT Protein
XX PN WO9918219-A1.
XX PD 15-APR-1999.
XX PF 03-OCT-1997; 97WO-US017864.
XX PR 03-OCT-1997; 97WO-US017864.
XX PA (PROC) PROCTER & GAMBLE CO.
XX PI Kitado H, Yoshikawa A, Zaiki T;
XX DR WPI; 1999-287737/24.
XX DR N-PSDB; AAX60578.
XX PT Novel human protease useful for treating or preventing skin flaking.
XX PS Claim 2; Page 32-34; 43pp; English.
XX CC This represents a human keratinocyte derived protease (KDP) polypeptide.
XX CC The protease is substantially similar to a sequence encoded by the insert
XX CC in plasmid PERM BP-6129. The KDP polypeptide, from human keratinocytes,
XX CC is specifically used in skin care products, particularly to treat or
XX CC prevent skin flaking, e.g. dandruff, and in laundry detergent
XX CC compositions. More generally the KDP polypeptide can be used in any
XX CC cleaning composition, e.g. hard surface or dishwashing products

Location/Qualifiers
1. 66
/note= "signal peptide"
67. 294
/note= "mature protein"

```

XX SQ Sequence 293 AA;
Query Match 100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Db 61 DSSSRRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLKLNRRIRPTKDVPRINVSCHPS 180
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLKLNRRIRPTKDVPRINVSCHPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

Qy 241 CQDSSGGPVVCGSLQGLVSGDYPFCARPNRPGVYTNLCCKFTKWIQETIQANS 293
Db 241 CQDSSGGPVVCGSLQGLVSGDYPFCARPNRPGVYTNLCCKFTKWIQETIQANS 293

RESULT 2
AAAY30524
ID AAAY30524 standard; protein; 293 AA.
XX AC AAAY30524;
XX DT 03-DEC-1999 (first entry)
XX DE Human PDSP-1 protein.
XX KW PDSP-1; Tango 114; prostate-derived serine protease; fusion protein;
XX KW treatment; proliferative disorder; prostate cancer; screening;
XX KW predictive medicine; growth factor; growth factor biosynthesis;
XX KW cellular proliferation; growth factor binding protein; metastasis;
XX KW cellular differentiation; prostate development; detection; forensic;
XX KW serine protease; human.
XX OS Homo sapiens.
XX FN WO9946391-A2.
XX PD 16-SEP-1999.
XX PF 12-MAR-1999; 99WO-US005416.
XX PR 12-MAR-1998; 98US-00041400.
XX PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX FI Holtzman DA;
XX DR WPI; 1999-561678/47.
XX DR N-PSDB; AAZ23317, AAZ23318.
XX PT Novel prostate derived serine protease polynucleotides and polypeptides
XX used to modulate cellular processes.
XX PS Claim 1c; Fig 1A; 94pp; English.
XX CC This invention describes a novel human prostate-derived serine protease
XX (PDSP) polynucleotide (also known as Tango 114). The prostate derived
XX serine protease (PDSP) polypeptide is used to modulate a variety of
XX cellular processes. It can be used to produce fusion proteins. PDSP is used
XX to treat proliferative disorders, e.g. prostate cancer. The protein may
XX also be used to produce antibodies, and to identify antagonists and

```

```

CC agonists. The PDSP polynucleotides, polypeptides, homologs and antibodies
CC can be used in screening assays; predictive medicine; and methods of
CC treatment. PDSP cleaves growth factors, and can be used for the
CC modulation of growth factor biosynthesis; generation of active peptides;
CC regulation of cellular proliferation; degradation of growth factor
CC binding proteins; regulation of cellular differentiation; regulation of
CC metastasis; and regulation of prostate development. The PDSP
CC polynucleotides can be used to express the protein; to detect PDSP mRNA;
CC to detect genetic alterations in the PDSP gene; in forensic biology; and
CC as a source of primers and probes. As serine protease have important
CC roles in cellular processes, there exists a need for identifying novel
CC serine protease such as the prostate derived serine protease (PDSP) of
CC the invention. This sequence represents the human PDSP-1 described in the
CC invention
XX SQ Sequence 293 AA;

```

```

Query Match 100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Db 61 DSSSRRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLKLNRRIRPTKDVPRINVSCHPS 180
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLKLNRRIRPTKDVPRINVSCHPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

Qy 241 CQDSSGGPVVCGSLQGLVSGDYPFCARPNRPGVYTNLCCKFTKWIQETIQANS 293
Db 241 CQDSSGGPVVCGSLQGLVSGDYPFCARPNRPGVYTNLCCKFTKWIQETIQANS 293

```

```

RESULT 3
AAAY38412
ID AAAY38412 standard; protein; 293 AA.
XX AC AAAY38412;
XX DT 30-SEP-1999 (first entry)
XX DE Human secreted protein encoded by gene No. 27.
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX KW developmental abnormality; foetal deficiency; blood; allergy; renal;
XX KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX FN WO9935158-A1.
XX PD 15-JUL-1999.
XX PF 06-JAN-1999; 99WO-US000108.
XX PR 07-JAN-1998; 98US-0070657P.
XX PR 07-JAN-1998; 98US-0070658P.
XX PR 07-JAN-1998; 98US-0070659P.
XX PR 07-JAN-1998; 98US-0070704P.

```



```

XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX Ruben SM, Soppet DR, Ebner R, Lafleur DW, Ni J, Brewer LA;
XX PI Olsen HS, Duan RD, Rosen CA;
XX PD
XX WPI; 1999-444190/37.
XX DR N-PSDB; AA206245.
XX PT
XX New isolated human genes and the secreted polypeptides they encode.
XX PS
XX Claim 11; Page 188-189; 227pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AA206210) for increasing the stability of the fused protein
XX as compared to the human protein only. The invention relates to 36 novel
XX genes and their fragments (nucleic acid sequences: AA206219-206263; amino
XX acid sequences AAY38386-Y38498) which are useful for preventing, treating
XX or ameliorating medical conditions e.g. by protein or gene therapy. Also,
XX pathological conditions can be diagnosed by determining the amount of the
XX new polypeptides in a sample or by determining the presence of mutations
XX in the new polynucleotides. Specific uses are described for each of the
XX 36 polynucleotides, based on which tissues they are most highly expressed
XX in (see AA206219 for described uses)
XX
XX Sequence 293 AA;
XX
Query Match          100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
DB 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
QY 61 DDSSRIINGSCDCHMTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
DB 61 DDSSRIINGSCDCHMTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
QY 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRIRTKDVRPINVSSHCP 180
DB 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRIRTKDVRPINVSSHCP 180
QY 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPGRPVTNLCCKTKWQIQTQANS 293
DB 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPGRPVTNLCCKTKWQIQTQANS 293
RESULT 4
AAY38426
ID AAY38426 standard; protein; 293 AA.
XX
XX AAY38426;
XX
XX 30-SEP-1999 (first entry)
XX
XX Human secreted protein encoded by gene No. 27.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; osteoarthritis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

```

```

OS Homo sapiens.
XX
XX WO9935158-A1.
XX
XX 15-JUL-1999.
XX
XX 06-JAN-1999; 99WO-US000108.
XX
XX 07-JAN-1998; 98US-0070657P.
XX
XX 07-JAN-1998; 98US-0070658P.
XX
XX 07-JAN-1998; 98US-0070692P.
XX
XX 07-JAN-1998; 98US-0070704P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Soppet DR, Ebner R, Lafleur DW, Ni J, Brewer LA;
XX PI Olsen HS, Duan RD, Rosen CA;
XX
XX WPI; 1999-444190/37.
XX
XX N-PSDB; AA206259.
XX
XX New isolated human genes and the secreted polypeptides they encode.
XX
XX Claim 11; Page 200-201; 227pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AA206210) for increasing the stability of the fused protein
XX as compared to the human protein only. The invention relates to 36 novel
XX genes and their fragments (nucleic acid sequences: AA206219-206263; amino
XX acid sequences AAY38386-Y38498) which are useful for preventing, treating
XX or ameliorating medical conditions e.g. by protein or gene therapy. Also,
XX pathological conditions can be diagnosed by determining the amount of the
XX new polypeptides in a sample or by determining the presence of mutations
XX in the new polynucleotides. Specific uses are described for each of the
XX 36 polynucleotides, based on which tissues they are most highly expressed
XX in (see AA206219 for described uses)
XX
XX Sequence 293 AA;
XX
Query Match          100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
DB 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
QY 61 DDSSRIINGSCDCHMTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
DB 61 DDSSRIINGSCDCHMTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
QY 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRIRTKDVRPINVSSHCP 180
DB 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRIRTKDVRPINVSSHCP 180
QY 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPGRPVTNLCCKTKWQIQTQANS 293
DB 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPGRPVTNLCCKTKWQIQTQANS 293
RESULT 5
AAY66726
ID AAY66726 standard; protein; 293 AA.
XX
XX AAY66726;
XX
XX 05-APR-2000 (first entry)

```

XX Membrane-bound protein PRO1132.
DE
XX
XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping.
KW
XX Homo sapiens.
OS
XX WO9963088-A2.
XX
XX PD 09-DEC-1999.
XX
XX PF 02-JUN-1999; 99WO-US012252.
XX
XX 02-JUN-1998; 98US-0087607P.
XX 02-JUN-1998; 98US-0087609P.
XX 02-JUN-1998; 98US-0087759P.
XX 03-JUN-1998; 98US-0087827P.
XX 04-JUN-1998; 98US-0088021P.
XX 04-JUN-1998; 98US-0088025P.
XX 04-JUN-1998; 98US-0088028P.
XX 04-JUN-1998; 98US-0088029P.
XX 04-JUN-1998; 98US-0088030P.
XX 04-JUN-1998; 98US-0088033P.
XX 04-JUN-1998; 98US-0088326P.
XX 05-JUN-1998; 98US-0088167P.
XX 05-JUN-1998; 98US-0088202P.
XX 05-JUN-1998; 98US-0088212P.
XX 05-JUN-1998; 98US-0088217P.
XX 09-JUN-1998; 98US-0088655P.
XX 10-JUN-1998; 98US-0088722P.
XX 10-JUN-1998; 98US-0088730P.
XX 10-JUN-1998; 98US-0088734P.
XX 10-JUN-1998; 98US-0088738P.
XX 10-JUN-1998; 98US-0088740P.
XX 10-JUN-1998; 98US-0088741P.
XX 10-JUN-1998; 98US-0088742P.
XX 10-JUN-1998; 98US-0088810P.
XX 10-JUN-1998; 98US-0088811P.
XX 10-JUN-1998; 98US-0088824P.
XX 10-JUN-1998; 98US-0088825P.
XX 10-JUN-1998; 98US-0088826P.
XX 11-JUN-1998; 98US-0088858P.
XX 11-JUN-1998; 98US-0088861P.
XX 11-JUN-1998; 98US-0088863P.
XX 11-JUN-1998; 98US-0088876P.
XX 12-JUN-1998; 98US-0089090P.
XX 12-JUN-1998; 98US-0089105P.
XX 16-JUN-1998; 98US-0089440P.
XX 16-JUN-1998; 98US-0089512P.
XX 16-JUN-1998; 98US-0089514P.
XX 17-JUN-1998; 98US-0089532P.
XX 17-JUN-1998; 98US-0089538P.
XX 17-JUN-1998; 98US-0089598P.
XX 17-JUN-1998; 98US-0089599P.
XX 17-JUN-1998; 98US-0089600P.
XX 17-JUN-1998; 98US-0089653P.
XX 18-JUN-1998; 98US-0089801P.
XX 18-JUN-1998; 98US-0089907P.
XX 18-JUN-1998; 98US-0089908P.
XX 19-JUN-1998; 98US-0089947P.
XX 19-JUN-1998; 98US-0089948P.
XX 19-JUN-1998; 98US-0089952P.
XX 22-JUN-1998; 98US-0090248P.
XX 22-JUN-1998; 98US-0090252P.
XX 22-JUN-1998; 98US-0090254P.
XX 23-JUN-1998; 98US-0090349P.
XX 23-JUN-1998; 98US-0090355P.
XX 24-JUN-1998; 98US-0090429P.
XX 24-JUN-1998; 98US-0090431P.
XX 24-JUN-1998; 98US-0090435P.
XX 24-JUN-1998; 98US-0090444P.
XX 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090538P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090557P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090691P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091358P.
PR 01-JUL-1998; 98US-0091360P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091519P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091633P.
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.
PR 07-JUL-1998; 98US-0091978P.
PR 07-JUL-1998; 98US-0091982P.
PR 09-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 20-JUL-1998; 98US-0093339P.
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095325P.
PR 10-AUG-1998; 98US-0095916P.
PR 10-AUG-1998; 98US-0095929P.
PR 10-AUG-1998; 98US-0096012P.
PR 11-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096146P.
PR 12-AUG-1998; 98US-0096329P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096894P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096950P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097218P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097951P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.

PR 26-AUG-1998; 98US-0098014P.
PR 31-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 12-JAN-1999; 99US-0115565P.
XX (GETH) GENENTECH INC.
PA Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX WPI; 2000-072883/06.
DR N-PSDB; AA265070.
XX Membrane-bound proteins and related nucleotide sequences.
XX Claim 12; Fig 226; 822pp; English.
XX The invention provides membrane-bound PRO polypeptides and
CC polynucleotides encoding them. The PRO sequences of the invention were
CC identified based on extracellular domain homology screening. The PRO
CC sequences have homology with proteins including LDL receptors, TIE
CC ligands and various enzymes. The membrane-bound proteins and receptor
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
CC immunoadhesins, for instance, can be used as therapeutic agents to block
CC receptor-ligand interactions. The membrane-bound proteins can also be
CC employed for screening of potential peptide or small molecule inhibitors
CC of the relevant receptor/ligand interaction. The PRO encoding sequences
CC are useful as hybridization probes, in chromosome and gene mapping and in
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences will
CC also be useful for the preparation of PRO polypeptides, especially by
CC recombinant techniques
XX Sequence 293 AA;
SQ
Query Match 100.0%; Score 1608; DB 3; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60
Db 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60
QY 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
Db 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
QY 121 YSLSPVYESGQOMFGQVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVRRPINVSSHCP 180
Db 121 YSLSPVYESGQOMFGQVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVRRPINVSSHCP 180
QY 181 AGTKCLVSGWGTTSKPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFACGDKAGRDS 240
Db 181 AGTKCLVSGWGTTSKPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFACGDKAGRDS 240
QY 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARPNGPVVYTNLCKFTKWIQETIQANS 293
Db 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARPNGPVVYTNLCKFTKWIQETIQANS 293
RESULT 6
AAB21296
ID AAB21296 standard; protein; 293 AA.
XX AAB21296;
AC AAB21296;
XX 02-FEB-2001 (first entry)
XX Human KLK-L2 protein.
XX Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KW kallikrein-like protein; serine protease; cytostatic; cancer;
KW prostrate cancer.
XX

OS Homo sapiens.
XX WO200053776-A2.
XX 14-SEP-2000.
XX 09-MAR-2000; 2000WO-CA000258.
XX 11-MAR-1999; 99US-0124260P.
PR 01-APR-1999; 99US-0127386P.
PR 21-JUL-1999; 99US-0144919P.
XX (MOUN) MOUNT SINAI HOSPITAL.
XX Yousef GM, Diamandis EP;
XX WPI; 2000-587440/55.
DR N-PSDB; AAA95905.
XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
PT protein mediated disorders, especially cancer.
XX Claim 9; Page 149-150; 184pp; English.
XX The present sequence is kallikrein-like protein KLK-L2. Kallikreins and
CC kallikrein-like proteins are a subgroup of the serine protease enzyme
CC family. They catalyze the selective cleavage of specific polypeptide
CC precursors to release peptides with potent biological activity. Nucleic
CC acids encoding kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4,
CC KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the
CC treatment, monitoring and diagnosis of cancers, especially prostate
CC cancer. They can also be used to identify a substance that can associate
CC with or mediate the biological activity of the proteins. Antibodies can
CC be used to treat conditions mediated by the kallikrein-like proteins
XX Sequence 293 AA;
SQ
Query Match 100.0%; Score 1608; DB 3; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60
Db 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60
QY 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
Db 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
QY 121 YSLSPVYESGQOMFGQVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVRRPINVSSHCP 180
Db 121 YSLSPVYESGQOMFGQVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVRRPINVSSHCP 180
QY 181 AGTKCLVSGWGTTSKPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFACGDKAGRDS 240
Db 181 AGTKCLVSGWGTTSKPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFACGDKAGRDS 240
QY 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARPNGPVVYTNLCKFTKWIQETIQANS 293
Db 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARPNGPVVYTNLCKFTKWIQETIQANS 293
RESULT 7
AAU12399
ID AAU12399 standard; protein; 293 AA.
XX AAU12399;
AC AAU12399;
XX 24-OCT-2001 (first entry)
XX Human PRO1132 polypeptide sequence.
XX Human secretory and transmembrane; PRO; mammalian; cancer; breast;
KW

PR 15-SEP-1999; 99WO-US021547.
 PR 08-OCT-1999; 99US-0158663P.
 PR 30-NOV-1999; 99WO-US028313.
 PR 01-DEC-1999; 99WO-US028301.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US0003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrarini N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Gurney AL, Kijavini IU, Napier MA, Pan J, Paoni NF;
 PI Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM, Wood WI,
 PI Zhang Z;
 XX
 XX WPI: 2001-032160/04.
 DR N-PSDB; AAF44216.
 XX
 PR PRO polynucleotides used to produce polypeptides used to target bioactive
 PT molecules such as toxins, radiolabels or antibodies, to specific cells,
 PT to cause targeted cell death.
 XX
 XX Claim 12; Fig 226; 935pp; English.
 XX
 CC The present invention describes human secreted and transmembrane PRO
 CC proteins. The PRO proteins have cytostatic activity. The PRO proteins can
 CC be used for targeted delivery of bioactive molecules, such as toxins,
 CC radiolabels or antibodies, that cause cell death. PRO nucleotide
 CC sequences, and their fragments, can be used as hybridisation probes, in
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA and
 CC DNA. They may also be used to produce transgenic animals which are used
 CC to develop and screen therapeutically useful reagents. The PRO nucleotide
 CC and protein sequence can be used for tissue typing and in treating
 CC cancer. Anti-PRO antibodies can be used in diagnostic assays. AAF44270 to
 CC AAF44470 represent PCR primers and hybridisation probes used in the
 CC isolation of human PRO sequences. AAF44087 to AAF44269 and AAF65154 to
 CC AAF65300 represent human PRO polynucleotide and protein sequences given
 CC in the exemplification of the present invention
 XX
 SQ Sequence 293 AA;
 Query Match 100.0%; Score 1608; DB 4; Length 293;
 Best Local Similarity 100.0%; Pred. No. 4.2e-120;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATATPPMMVLCALITALLGVTEHVLANDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60
 DB 1 MATATPPMMVLCALITALLGVTEHVLANDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60
 QY 61 DSSSRITNGSDCMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
 DB 61 DSSSRITNGSDCMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
 QY 121 YLSLSPVYESGQMGFGVKSIPHPGYSHPGHNDMLKLNRIIRTKDVRPINVSSHCPSS 180
 DB 121 YLSLSPVYESGQMGFGVKSIPHPGYSHPGHNDMLKLNRIIRTKDVRPINVSSHCPSS 180
 QY 181 AGTKCLVSGWGTTSKPQVHPFKVLCQLNISVLISQKRCEDAYPRQIDDTMFCAKAGKGRDS 240
 DB 181 AGTKCLVSGWGTTSKPQVHPFKVLCQLNISVLISQKRCEDAYPRQIDDTMFCAKAGKGRDS 240
 QY 241 CQDGGGPPVNCVGLSVNGDYPGPNRPGVYTNLCCKTKWQIETIQANS 293
 |||||

Db 241 CQDGGGPPVNCVGLSVNGDYPGPNRPGVYTNLCCKTKWQIETIQANS 293
 RESULT 9
 AAU81966
 ID AAU81966 standard; protein; 293 AA.
 XX
 AC AAU81966;
 XX
 DT 09-APR-2002 (first entry)
 DE Human PRO1132.
 XX
 KW Human; PRO: antiinflammatory; ophthalmological; vasotropic;
 KW retinal cell injury; ocular disease; retinitis pigmentosa;
 KW macular degeneration; retinal detachment; retinal tear; retinopathy;
 KW retinal degenerative disease; macular hole; degenerative myopia;
 KW acute retinal necrosis syndrome; traumatic chorioretinopathy;
 KW Purtscher's retinopathy; oedema; ischaemic condition;
 KW retinal vision occlusion; collagen vascular disease;
 KW thrombocytopenic purpura; uveitis; retinal vasculitis; Eales disease;
 KW systemic lupus erythematosus; environmental trauma.
 XX
 OS Homo sapiens.
 XX
 PN WO200109327-A2.
 XX
 PD 08-FEB-2001.
 XX
 PF 28-JUL-2000; 2000WO-US020710.
 XX
 PR 28-JUL-1999; 99US-0146222P.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 29-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 01-DEC-1999; 99WO-US028301.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Goddard A, Godowski PJ, Gurney AL;
 PI Kijavini IU, Lafleur M, Mark MR, Marsters SA, Pitti RM, Watanabe CK;
 PI Wood WI;
 XX
 DR WPI: 2002-130120/17.
 DR N-PSDB; ABK28605.
 XX
 PT Promoting survival of retinal cells, or delaying or preventing retinal
 PT cell injury or death, by contacting retinal cells with PRO175, 220, 216,
 PT 243, 306, 346, 322, 536, 943, 840, 828, 826, 1068 or PRO1132 polypeptide.
 XX
 PS Claim 44; Fig 29; 152pp; English.
 XX
 CC The invention relates to promoting the survival of retinal cells, or
 CC delaying or preventing retinal cell injury or death, by contacting the
 CC retinal cells with the polypeptide such as PRO175, PRO220, PRO216,
 CC PRO243, PRO306, PRO346, PRO322, PRO336, PRO943, PRO840, PRO828, PRO826,
 CC PRO1068 or PRO1132 polypeptide. Also included are the nucleic acids
 CC encoding the PRO proteins, a vector comprising the nucleic acid, a host
 CC cell comprising the vector, and anti-PRO antibody. The PRO proteins are
 CC useful for promoting survival of retinal cells (retinal neurons such as
 CC retinal ganglion cells, displaced retinal ganglion cells, amacrine cells,
 CC displaced amacrine cells, horizontal neurons or bipolar neurons, rod

CC	photoreceptors, or supportive cells such as Muller cells or pigment epithelial cells), or delaying or preventing retinal cell injury or death caused by ocular disease (which is or is associated with retinitis pigmentosa, macular degeneration, retinal detachment, retinal tear, retinopathy, retinal degenerative disease, macular hole, degenerative myopia, acute retinal necrosis syndrome, traumatic choriorretinopathy or contusion, Purtscher's retinopathy, oedema, an ischaemic condition, central or peripheral retinal vision occlusion, collagen vascular disease, thrombocytopenic purpura, uveitis, retinal vasculitis, occlusion CC associated with Eales disease or systemic lupus erythematosus), retinal injury or environmental trauma. The retinal cell injury or death is delayed or prevented by substantially not causing angiogenesis or CC mitogenesis. The present sequence represents a PRO protein
xx	
SQ	Sequence 293 AA;
	Query Match 100.0%; Score 1608; DB 5; Length 293;
	Best Local Similarity 100.0%; Pred. No. 4.2e-120;
	Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 MATARPPMMWVLCALITALLLGVTGTEHVLANNDSVDHPSTNTVPSSNQDLGAGAGEDARS 60
Dd	1 MATARPPMMWVLCALITALLLGVTGTEHVLANNDSVDHPSTNTVPSSNQDLGAGAGEDARS 60
Qy	61 DSSSRRIINGSDCMHTQPQAALLRPNLQYCGAVLVHPQLLTAACHCKKVFVRVLGH 120
Dd	61 DSSSRRIINGSDCMHTQPQAALLRPNLQYCGAVLVHPQLLTAACHCKKVFVRVLGH 120
Qy	121 YLSFVVYESGOOMFGVKSIIPHPGYSHPGHSNDMLIKLNRRIRPTKDVRPINVSSHCPFS 180
Dd	121 YLSFVVYESGOOMFGVKSIIPHPGYSHPGHSNDMLIKLNRRIRPTKDVRPINVSSHCPFS 180
Qy	181 AGTKCLVSGWGTTKSPQVHFPAKLQCLINISVLQKRCEADYPRQIDDTMFCAGDKAGRDS 240
Dd	181 AGTKCLVSGWGTTKSPQVHFPAKLQCLINISVLQKRCEADYPRQIDDTMFCAGDKAGRDS 240
Qy	241 CQGDSGGPVCNGSLQGLVSWGDPYCARPNRPGVYTNLCKFTKWIQETIQANS 293
Dd	241 CQGDSGGPVCNGSLQGLVSWGDPYCARPNRPGVYTNLCKFTKWIQETIQANS 293
RESULT 10	
ABU58064	ID ABU58064 standard; protein; 293 AA.
XX AC	ABU58064;
XX DT	14-APR-2003 (first entry)
XX DE	Human PRO polypeptide #96.
XX KW	Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
XX KW	horse; cow; dog; cat; sheep; pig; goat; rabbit; ADEPT;
XX KW	antibody-dependent enzyme mediated prodrug therapy.
XX OS	Homo sapiens.
XX PN	US2003027163-A1.
XX PD	06-FEB-2003.
XX PF	15-NOV-2001; 2001US-00997666.
XX PR	16-JUN-1997; 97US-0049787P.
XX PR	17-OCT-1997; 97US-0062250P.
XX PR	05-NOV-1997; 97WO-USO20069.
XX PR	12-NOV-1997; 97US-0065186P.
XX PR	13-NOV-1997; 97US-0065311P.
XX PR	24-NOV-1997; 97US-0066770P.
XX PR	25-FEB-1998; 98US-0075943P.
XX PR	20-MAR-1998; 98US-0078910P.
XX PR	28-APR-1998; 98US-0083322P.
XX PR	07-MAY-1998; 98US-0084600P.

```
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.
PR 07-JUL-1998; 98US-0091978P.
PR 07-JUL-1998; 98US-0091982P.
PR 09-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 20-JUL-1998; 98US-0093333P.
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095322P.
PR 10-AUG-1998; 98US-0095916P.
PR 10-AUG-1998; 98US-0095922P.
PR 10-AUG-1998; 98US-0096012P.
PR 11-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096146P.
PR 12-AUG-1998; 98US-0096323P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096894P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096950P.
PR 18-AUG-1998; 98US-0096953P.
PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097219P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 31-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100858P.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 22-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 12-MAR-1999; 99US-0123957P.
PR 02-JUN-1999; 99WO-US012252.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0143048P.
PR 26-JUL-1999; 99US-0144758P.
PR 28-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 17-AUG-1999; 99US-0149336P.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 30-OCT-1999; 99US-0158663P.
PR 08-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-JUN-2000; 2000US-0213637P.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 07-SEP-2000; 2000US-0230978P.

Query Match 100.0%; Score 1608; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMVLCALITALLLGVTGHEVLANNVSCDHPSTNTPVSGSNODLGAGAGEDARS 60
DB 1 MATARPPMMVLCALITALLLGVTGHEVLANNVSCDHPSTNTPVSGSNODLGAGAGEDARS 60
QY 61 DSSSRIINGSDCDMHTQPMQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120
DB 61 DSSSRIINGSDCDMHTQPMQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120
QY 121 YLSLPTVESGQMFQGVKSIHPGYSHPGHNDLMLKLNRRIRPTKDVRIINVSSHCP 180
DB 121 YLSLPTVESGQMFQGVKSIHPGYSHPGHNDLMLKLNRRIRPTKDVRIINVSSHCP 180
QY 181 AGTKCLVSGMGTTSKSPQVHPFKVQLCLNISVLISOKRCEDAYPROIDDTMFAGKAGRDS 240
DB 181 AGTKCLVSGMGTTSKSPQVHPFKVQLCLNISVLISOKRCEDAYPROIDDTMFAGKAGRDS 240
QY 241 CQGDGGPVVVCNGLSLQGLVSWGDYPCARPNNRPGVYTNLCKFTKWIQETIQANS 293
DB 241 CQGDGGPVVVCNGLSLQGLVSWGDYPCARPNNRPGVYTNLCKFTKWIQETIQANS 293

RESULT 11
ABU59142
ID ABU59142 standard; protein; 293 AA.
XX
AC ABU59142;
XX
DT 28-APR-2003 (first entry)
XX
DE Novel human secreted or transmembrane protein PRO1132.
XX
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumor; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disorder;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW chondrocyte redifferentiation; sports injury; arthritis.
XX
OS Homo sapiens.
XX
```

PN US2002132252-A1.
XX 19-SEP-2002.
XX 14-NOV-2001; 2001US-00990442.
PR 16-JUN-1997; 97US-0049787P.
PR 17-OCT-1997; 97US-0062257P.
PR 05-NOV-1997; 97WO-US020069.
PR 12-NOV-1997; 97US-0065186P.
PR 13-NOV-1997; 97US-0065311P.
PR 24-NOV-1997; 97US-0066770P.
PR 25-FEB-1998; 98US-0075945P.
PR 20-MAR-1998; 98US-0078910P.
PR 28-APR-1998; 98US-0083322P.
PR 07-MAY-1998; 98US-0084600P.
PR 28-MAY-1998; 98US-0087106P.
PR 02-JUN-1998; 98US-0087607P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088021P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088026P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088030P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088859P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089801P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 98WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 02-JUN-1999; 99WO-US012252.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 06-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 28-AUG-2001; 2001US-00941992.
XX (GETH) GENENTECH INC.
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NP;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams EM, Wood WI;
PI Zhang Z;
XX WPI; 2003-2477083/24.
DR N-PSDB; ABX80317.
XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
PT are therapeutically useful for enhancing immune response and in cancer
PT treatments.
XX Claim 12; Fig 226; 648pp; English.
XX The invention describes an isolated human PRO polypeptide. The PRO
CC polypeptides are useful in detecting PRO polypeptides in a sample, in
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and
CC in modulating at least one biological activity of a cell expressing a PRO
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186
CC stimulate adrenal cortical capillary endothelial growth and PRO536,
CC PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO126,
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus
CC useful for treating conditions or disorders where angiogenesis would be
CC beneficial, e.g. wound healing and antagonist of this polypeptide are
CC useful for treating cancerous tumours. PRO812 inhibits vascular
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial
CC cells and is thus useful for inhibiting endothelial cell growth in
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,
CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of
CC stimulated T-lymphocytes and are therapeutically useful for enhancing
CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of
CC rod photoreceptor cells) and therefore are useful for treating retinal
CC disorders of injuries, e.g. retinitis pigmentosum, AMD. PRO819, PRO813
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,
CC and therefore are useful for treating kidney disorders associated with
CC decreased mesangial cell function such as Berger disease or Crohn's
CC nephropathies associated with dermatitis, herpeticiformis or Crohn's
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
CC proliferation and/or redifferentiation of chondrocytes in culture and are
CC thus useful for treating sports injuries, and arthritis. This is the

CC amino acid sequence of a novel human PRO protein
XX SQ Sequence 293 AA;

Query Match 100.0%; Score 1608; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPSCNODLGAGAGEDARS 60
DB 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPSCNODLGAGAGEDARS 60
QY 61 DSSSRRIINGSDCMHTOPWQAALLLRPNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120
DB 61 DSSSRRIINGSDCMHTOPWQAALLLRPNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120
QY 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNDMLIKLNRRIRPTKQVRPINVSSHCP 180
DB 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNDMLIKLNRRIRPTKQVRPINVSSHCP 180
QY 181 AGTKLVSGWGTTSQVHPKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKLVSGWGTTSQVHPKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQDSGGPVVNCVSLQGLVSGWGYPCARPNGVYTNLCCKTKWIQETIQANS 293
DB 241 CQDSGGPVVNCVSLQGLVSGWGYPCARPNGVYTNLCCKTKWIQETIQANS 293

RESULT 12
ABU82654
ID ABU82654 standard; protein; 293 AA.

XX AC ABU82654;

XX DT 26-JUN-2003 (first entry)

XX DE Human secreted/transmembrane protein PRO1132.

XX KW Human; PRO; secreted protein; transmembrane protein;
cardiac insufficiency disorders; angiogenesis; wound healing;
cancerous tumour; immune response; retinal disorder; sight loss;
retinitis pigmentosum; age-related macular degeneration; AMD;
kidney disorder; Berger disease; nephropathy; dermatitis; herpeticiformis;
Crohn's disease; sports injury; arthritis.

XX OS Homo sapiens.

XX PN US2003032023-A1.

XX PD 13-FEB-2003.

XX PF 14-NOV-2001; 2001US-00990711.

XX PR 16-JUN-1997; 97US-0049787P.

PR 17-OCT-1997; 97US-0062250P.

PR 05-NOV-1997; 97WO-US020069.

PR 12-NOV-1997; 97US-0065186P.

PR 13-NOV-1997; 97US-0065311P.

PR 24-NOV-1997; 97US-0066770P.

PR 25-FEB-1998; 98US-0075945P.

PR 20-MAR-1998; 98US-0078910P.

PR 28-APR-1998; 98US-0083322P.

PR 07-MAY-1998; 98US-0084600P.

PR 28-MAY-1998; 98US-0087106P.

PR 02-JUN-1998; 98US-0087607P.

PR 02-JUN-1998; 98US-0087609P.

PR 02-JUN-1998; 98US-0087759P.

PR 03-JUN-1998; 98US-0087827P.

PR 04-JUN-1998; 98US-0088021P.

PR 04-JUN-1998; 98US-0088025P.

PR 04-JUN-1998; 98US-0088026P.

PR 04-JUN-1998; 98US-0088028P.

PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088030P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088856P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089801P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089947P.
PR 19-JUN-1998; 98US-0089948P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 23-JUN-1998; 98US-0090349P.
PR 23-JUN-1998; 98US-0090355P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090431P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090542P.
PR 24-JUN-1998; 98US-0090557P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091360P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091519P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091633P.
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.
PR 07-JUL-1998; 98US-0091978P.
PR 07-JUL-1998; 98US-0091982P.
PR 09-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 20-JUL-1998; 98US-0093339P.
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.

PR	14-SEP-1998;	98WO-US019094.	PR	10-MAY-2001;	2001US-00854280.
PR	14-SEP-1998;	98WO-US019177.	PR	18-MAY-2001;	2001US-00860216.
PR	16-SEP-1998;	98WO-US019330.	PR	25-MAY-2001;	2001US-00866028.
PR	17-SEP-1998;	98WO-US019437.	PR	25-MAY-2001;	2001US-00866034.
PR	07-OCT-1998;	98WO-US021141.	PR	01-JUN-2001;	2001US-00872035.
PR	29-OCT-1998;	98WO-US022991.	PR	01-JUN-2001;	2001US-00872035.
PR	29-OCT-1998;	98WO-US022992.	PR	05-JUN-2001;	2001US-00874503.
PR	20-NOV-1998;	98WO-US024855.	PR	14-JUN-2001;	2001US-00882636.
PR	01-DEC-1998;	98WO-US025108.	PR	19-JUN-2001;	2001US-00886342.
PR	05-JAN-1999;	98WO-US000106.	PR	20-JUN-2001;	2001US-00886342.
PR	08-MAR-1999;	98WO-US005028.	PR	21-JUN-2001;	2001US-00887879.
PR	10-MAR-1999;	98WO-US00505190.	PR	22-JUN-2001;	2001US-00887879.
PR	20-APR-1999;	98WO-US008615.	PR	29-JUN-2001;	2001US-00908827.
PR	14-MAY-1999;	98WO-US010733.	PR	09-JUL-2001;	2001US-00921735.
PR	02-JUN-1999;	98WO-US012252.	PR	18-JUL-2001;	2001US-00924419.
PR	01-SEP-1999;	98WO-US020111.	PR	06-AUG-2001;	2001US-00927796.
PR	08-SEP-1999;	98WO-US020594.	PR	16-AUG-2001;	2001US-00931836.
PR	13-SEP-1999;	98WO-US020944.	XX	19-DEC-2001;	2001US-00028072.
PR	15-SEP-1999;	98WO-US021090.	PA	(GETH) GENENTECH INC.	
PR	15-SEP-1999;	98WO-US021547.	XX	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;	
PR	05-OCT-1999;	98WO-US023089.	PI	Gerriksen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;	
PR	29-NOV-1999;	98WO-US028214.	PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;	
PR	30-NOV-1999;	98WO-US028313.	XX	WPI; 2003-341980/32.	
PR	30-NOV-1999;	98WO-US028409.	DR	N-PSDB; ACD24080.	
PR	01-DEC-1999;	98WO-US028301.	XX	New secreted and transmembrane PRO nucleic acids, for treating	
PR	01-DEC-1999;	98WO-US028634.	PT	inflammation, organ failure, atherosclerosis, cardiac injury,	
PR	02-DEC-1999;	98WO-US028551.	PT	infertility, birth defects, premature aging, acquired immunodeficiency	
PR	02-DEC-1999;	98WO-US028584.	PT	syndrome (AIDS), or cancer.	
PR	02-DEC-1999;	98WO-US028585.	XX	Claim 12; Fig 456; 660pp; English.	
PR	16-DEC-1999;	98WO-US030095.	PS	The invention describes an isolated nucleic acid (I) comprising, or which	
PR	20-DEC-1999;	98WO-US030911.	XX	has 80 % sequence identity to, or the full-length coding sequence of, one	
PR	20-DEC-1999;	98WO-US030999.	CC	of 275 nucleotide sequences, and which encodes a corresponding	
PR	22-DEC-1999;	98WO-US030720.	CC	polypeptide selected from 275 amino acid sequences, where all sequences	
PR	30-DEC-1999;	98WO-US031243.	CC	are given in the specification. The polypeptide encoded by (I) is used to	
PR	30-DEC-1999;	98WO-US031274.	CC	detect PRO polypeptides, link a bioactive molecule to a cell expressing a	
PR	05-JAN-2000;	2000WO-US000219.	CC	PRO polypeptide, modulate a biological activity of a cell, stimulate the	
PR	06-JAN-2000;	2000WO-US000277.	CC	release of tumour necrosis factor (TNF)-alpha from human blood, modulate	
PR	11-FEB-2000;	2000WO-US000376.	CC	the uptake of glucose or free fatty acid by cells, stimulate or inhibit	
PR	18-FEB-2000;	2000WO-US004341.	CC	the proliferation or differentiation of cells or gene expression,	
PR	18-FEB-2000;	2000WO-US004342.	CC	stimulate the release of proteoglycans, stimulate the release of cytokine	
PR	22-FEB-2000;	2000WO-US004414.	CC	from peripheral blood mononuclear cells, inhibit the binding of A-peptide	
PR	24-FEB-2000;	2000WO-US004914.	CC	to factor VIIa, or detect the presence of tumour in a mammal. The nucleic	
PR	01-MAR-2000;	2000WO-US005004.	CC	acid and polypeptide encoded by it, are useful for treating inflammatory	
PR	02-MAR-2000;	2000WO-US00505601.	CC	diseases, organ failure, atherosclerosis, cardiac injury, infertility,	
PR	02-MAR-2000;	2000WO-US005746.	CC	birth defects, premature aging, acquired immunodeficiency syndrome	
PR	10-MAR-2000;	2000WO-US005841.	CC	(AIDS), cancer, or diabetic complications. The nucleic acid is useful as	
PR	15-MAR-2000;	2000WO-US006319.	CC	hybridisation probes, in chromosome and gene mapping, and in generating	
PR	20-MAR-2000;	2000WO-US006884.	CC	antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,	
PR	21-MAR-2000;	2000WO-US007377.	CC	diagnostics, biosensors or bioreactors. Both are useful in tissue typing.	
PR	30-MAR-2000;	2000WO-US007532.	CC	This is the amino acid sequence of a novel human secreted and	
PR	17-MAY-2000;	2000WO-US013705.	CC	transmembrane PRO polypeptide	
PR	22-MAY-2000;	2000WO-US014042.	XX	Sequence 293 AA;	
PR	30-MAY-2000;	2000WO-US014941.	SQ	Query Match 100.0%; Score 1608; DB 6; Length 293;	
PR	02-JUN-2000;	2000WO-US015264.		Best Local Similarity 100.0%; Pred. No. 4.2e-120;	
PR	28-JUL-2000;	2000WO-US020710.		Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
PR	11-AUG-2000;	2000WO-US022031.		1 MATARPPMMVLCALITALLGLGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGEDARS 60	
PR	23-AUG-2000;	2000WO-US023522.		1 MATARPPMMVLCALITALLGLGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGEDARS 60	
PR	24-AUG-2000;	2000WO-US023328.		61 DSSSRIINGSDCDMHTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120	
PR	08-NOV-2000;	2000WO-US030952.		61 DSSSRIINGSDCDMHTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120	
PR	10-NOV-2000;	2000WO-US030873.			
PR	01-DEC-2000;	2000WO-US032678.			
PR	20-DEC-2000;	2000US-0074259.			
PR	20-DEC-2000;	2000WO-US034956.			
PR	28-FEB-2001;	2001US-00796498.			
PR	28-FEB-2001;	2001WO-US006520.			
PR	01-MAR-2001;	2001WO-US006666.			
PR	09-MAR-2001;	2001US-00802706.			
PR	14-MAR-2001;	2001US-00808689.			
PR	22-MAR-2001;	2001US-00816744.			
PR	05-APR-2001;	2001US-00828366.			
PR	10-MAY-2001;	2001US-00854208.			

```
Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVRRPINVSSHCPs 180
    |||||
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVRRPINVSSHCPs 180
    |||||
Qy 181 AGTKCLVSGWGTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDGS 240
    |||||
Db 181 AGTKCLVSGWGTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDGS 240
    |||||
Qy 241 CQDGGGPPVVCNGSLQGLVSWGDPFCARPNRPVYTNLCKFTKWIQETIQANS 293
    |||||
Db 241 CQDGGGPPVVCNGSLQGLVSWGDPFCARPNRPVYTNLCKFTKWIQETIQANS 293
    |||||

RESULT 14
ADA57425
ID ADA57425 standard; protein; 293 AA.
XX
AC ADA57425;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein #257.
XX
KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
KW cyostatic; cerebroprotective; neuroprotective; nootropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology.
XX
OS Homo sapiens.
XX
PN WO2002102994-A2.
XX
PD 27-DEC-2002.
XX
PF 19-MAR-2002; 2002WO-US008278.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI, 2003-167512/16.
DR N-FSDB; ADA56532.
XX

New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.
XX
FS Claim 13; SEQ ID NO 1618; 1754pp; English.
XX

The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 95% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that binds to
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
```

```
CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to one of the polypeptide of the invention. Note: The sequence data for
CC this patent did form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 293 AA;
```

```
Query Match 100.0%; Score 1608; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
    |||||
Db 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
    |||||
Qy 61 DDSSSRINSGDCDMHTQPHQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120
    |||||
Db 61 DDSSSRINSGDCDMHTQPHQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120
    |||||
Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVRRPINVSSHCPs 180
    |||||
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVRRPINVSSHCPs 180
    |||||
Qy 181 AGTKCLVSGWGTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDGS 240
    |||||
Db 181 AGTKCLVSGWGTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDGS 240
    |||||
Qy 241 CQDGGGPPVVCNGSLQGLVSWGDPFCARPNRPVYTNLCKFTKWIQETIQANS 293
    |||||
Db 241 CQDGGGPPVVCNGSLQGLVSWGDPFCARPNRPVYTNLCKFTKWIQETIQANS 293
    |||||
```

RESULT 15

ADA56974
ID ADA56974 standard; protein; 293 AA.

XX
AC ADA56974;

XX
DT 20-NOV-2003 (first entry)

XX
DE Human secreted protein #257.

XX immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
KW cyostatic; cerebroprotective; neuroprotective; nootropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology.

XX Homo sapiens.

OS WO2002102994-A2.

XX 27-DEC-2002.

PF 19-MAR-2002; 2002WO-US008278.

XX 21-MAR-2001; 2001US-0277340P.

Search completed: July 4, 2005, 15:21:15
Job time : 77 secs

PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-167512/16.
XX N-PSDB; ADA56078.
XX
XX New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.
XX
XX
PS Claim 13; SEQ ID NO 1164; 1754pp; English.
XX
XX The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 93% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that binds to
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
CC disorders (e.g. Parkinson's disease or Alzheimer's disease) and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to one of the polypeptide of the invention. Note: The sequence data for
CC this patent did form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 293 AA;
Query Match 100.0%; Score 1608; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MATARPPMMWVLCALITALLGVTEHVLANNVSCDHPSTNTVPSSNODLGAGAGEDARS 60
Db 1 MATARPPMMWVLCALITALLGVTEHVLANNVSCDHPSTNTVPSSNODLGAGAGEDARS 60
Qy 61 DSSSRIINGSDCDMHTQPWQAALLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH 120
Db 61 DSSSRIINGSDCDMHTQPWQAALLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH 120
Qy 121 YLSLSPVYESGQMGFGVKSIHPHGYSHPGHNDLMLIKLNRIRPTKDVRIINVSSHCPS 180
Db 121 YLSLSPVYESGQMGFGVKSIHPHGYSHPGHNDLMLIKLNRIRPTKDVRIINVSSHCPS 180
Qy 181 AGTKCLVSGWGTTSQVHFPKVLQCLNIVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTSQVHFPKVLQCLNIVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Qy 241 CQGSDDGGPVVNCNGSLQGLVSWGDI PCARPNPVGVYTNLCKFTKWIQETIQANS 293
Db 241 CQGSDDGGPVVNCNGSLQGLVSWGDI PCARPNPVGVYTNLCKFTKWIQETIQANS 293

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 01:15:51 ; Search time 24 Seconds
(without alignments)
911.341 Million cell updates/sec

Title: US-09-936-271C-14
Perfect score: 1608
Sequence: 1 MATARPPMMWVLCALITALL.....VYTNLCRFTKWIQETIQANS 293

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	293	4	US-09-509-908-2
2	1479	92.0	270	4	US-09-949-016-7712
3	1173	72.9	268	2	US-08-824-874-1
4	1173	72.9	268	3	US-09-210-084-1
5	1173	72.9	268	4	US-09-764-762-1
6	671	41.7	254	3	US-09-439-313-525
7	671	41.7	254	4	US-09-636-215-525
8	671	41.7	254	4	US-09-685-166A-525
9	671	41.7	254	4	US-09-679-426-525
10	671	41.7	254	4	US-09-759-143-525
11	671	41.7	254	4	US-09-651-236-525
12	666	41.4	254	3	US-09-439-313-523
13	666	41.4	254	4	US-09-636-215-523
14	666	41.4	254	4	US-09-685-166A-523
15	666	41.4	254	4	US-09-679-426-523
16	666	41.4	254	4	US-09-759-143-523
17	666	41.4	254	4	US-09-651-236-523
18	654.5	40.7	289	4	US-09-386-642-14
19	650.5	40.5	250	4	US-09-205-258-427
20	650.5	40.5	282	3	US-09-025-059-1
21	641.5	39.9	248	3	US-08-944-483-24
22	641.5	39.9	260	4	US-09-618-259-7
23	640	39.8	449	4	US-09-636-215-617
24	640	39.8	449	4	US-09-685-166A-617
25	640	39.8	449	4	US-09-679-426-617
26	640	39.8	449	4	US-09-759-143-617
27	640	39.8	449	4	US-09-651-236-617

28	639.5	39.8	260	3	US-09-008-271A-7	Sequence 7, Appli
29	638.5	39.7	260	3	US-09-070-526-2	Sequence 2, Appli
30	638.5	39.7	290	4	US-09-949-016-8166	Sequence 8166, Ap
31	638	39.7	248	1	US-08-744-026-1	Sequence 1, Appli
32	638	39.7	248	2	US-09-102-732-1	Sequence 1, Appli
33	638	39.7	248	3	US-09-261-767-1	Sequence 1, Appli
34	635	39.5	284	4	US-09-386-642-54	Sequence 54, Appli
35	632.5	39.3	288	4	US-09-386-642-13	Sequence 13, Appli
36	627	39.0	260	3	US-09-025-059-3	Sequence 3, Appli
37	627	39.0	260	4	US-09-618-259-8	Sequence 8, Appli
38	620	38.6	220	3	US-09-439-313-327	Sequence 327, App
39	620	38.6	220	3	US-09-352-616A-327	Sequence 327, App
40	620	38.6	220	4	US-09-232-149A-327	Sequence 327, App
41	620	38.6	220	4	US-09-636-215-327	Sequence 327, App
42	620	38.6	220	4	US-09-685-166A-327	Sequence 327, App
43	620	38.6	220	4	US-09-688-489-327	Sequence 327, App
44	620	38.6	220	4	US-09-679-426-327	Sequence 327, App
45	620	38.6	220	4	US-09-759-143-327	Sequence 327, App

ALIGNMENTS

RESULT 1
US-09-509-908-2
; Sequence 2, Application US/09509908
; Patent No. 6589770
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company, N/A N/A
; TITLE OF INVENTION: A protease
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: T. David Reed
; STREET: 5299 Spring Grove Avenue
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45217-1087
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/509,908
; FILING DATE: 28-Feb-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, T David
; REGISTRATION NUMBER: 32,931
; REFERENCE/DOCKET NUMBER: AA-264F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-627-7025
; TELEFAX: 513-627-6333
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-509-908-2

Query Match	100.0%;	Score 1608;	DB 4;	Length 293;
Best Local Similarity	100.0%;	Pred. No. 6e-163;		
Matches 293;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNODLGAGAGEDARS	60	
Db	1	MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNODLGAGAGEDARS	60	
Qy	61	DDSSRLINGSDDCMHTQPWQAALLLPNOLYCCAVLHPQWLLTAHCRKKVFRVLGH	120	

Db 61 DSSSRINGSDCDMHTQWQAALLLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSSHCPS 180
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSSHCPS 180
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQKRCEDAYPRQIDDTMFACGDKGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQKRCEDAYPRQIDDTMFACGDKGRDS 240
Qy 241 CQSDSGPVVWVNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
Db 241 CQSDSGPVVWVNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293

RESULT 2
US-09-949-016-7712
; Sequence 7712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7712
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7712

Query Match 92.0%; Score 1479; DB 4; Length 270;
Best Local Similarity 100.0%; Pred. No. 3e-149;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 EHVLANNDVSCDHPSTNTPVSGNODLGAGAGEDARSDSSRIINGSDCDMHTQWQAAL 84
Db 2 EHVLANNDVSCDHPSTNTPVSGNODLGAGAGEDARSDSSRIINGSDCDMHTQWQAAL 61
Qy 85 LLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPHG 144
Db 62 LLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPHG 121
Qy 145 YSHPGHNDMLIKLNRIRPTKDVRIINVSSHCPSAGTKCLVSGWGTTKSPQVHPFKVL 204
Db 122 YSHPGHNDMLIKLNRIRPTKDVRIINVSSHCPSAGTKCLVSGWGTTKSPQVHPFKVL 181
Qy 205 QCLNISVLSQKRCEDAYPRQIDDTMFACGDKAGRDSGQSDSGPVVWVNGSLQGLVSWGDY 264
Db 182 QCLNISVLSQKRCEDAYPRQIDDTMFACGDKAGRDSGQSDSGPVVWVNGSLQGLVSWGDY 241
Qy 265 PCARPNRPGVYTNLCKFTKWIQETIQANS 293
Db 242 PCARPNRPGVYTNLCKFTKWIQETIQANS 270

RESULT 3
US-08-824-874-1
; Sequence 1, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN

; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KERANOT02
; CLONE: 820694
US-08-824-874-1

Query Match 72.9%; Score 1173; DB 2; Length 268;
Best Local Similarity 85.7%; Pred. No. 1.2e-116;
Matches 221; Conservative 8; Mismatches 25; Indels 4; Gaps 2;

Qy 38 PSNTVPVSGSNQ--DLGAGAGEDARSDSSRIINGSDCDMHTQWQAALLRPNQLYCGA 95
Db 13 PLTPCPLGATRTWELGPKTPGRMTAAASSMDPTAICT--PSRQAALLRPNQLYCGA 70
Qy 96 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPGYSHPGHNDLM 155
Db 71 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPGYSHPGHNDLM 130
Qy 156 LILNRRIRPTKDVRIINVSSHCPSAGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQK 215
Db 131 LILNRRIRPTKDVRIINVSSHCPSAGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQK 190
Qy 216 RCEDAYPRQIDDTMFACGDKAGRDSGQSDSGPVVWVNGSLQGLVSWGDYPCARPNRPGVY 275
Db 191 RCEDAYPRQIDDTMFACGDKAGRDSGQSDSGPVVWVNGSLQGLVSWGDYPCARPNRPGVY 250
Qy 276 TNLCKFTKWIQETIQANS 293
Db 251 TNLCKFTKWIQETIQANS 268

RESULT 4
US-09-210-084-1
; Sequence 1, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,084
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/824,874
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0252 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 268 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: KERANOT02
 CLONE: 820694
 US-09-210-084-1

Query Match 72.9%; Score 1173; DB 3; Length 268;
 Best Local Similarity 85.7%; Pred. No. 1.2e-116;
 Matches 221; Conservative 8; Mismatches 25; Indels 4; Gaps 2;
 QY 38 PSNTVPSGNSQ--DLGAGAGEDARDSDSSRIINGSDCDMHTQPMQAAALLLRPNOLYCGA 95
 DB 13 PLTPCPLGATRTWELGPKTFCGRMTAAASMSMDPTAICT--PSRQQAALLLRPNOLYCGA 70
 QY 96 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIPIHPGYSHPGHNSDLM 155
 DB 71 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIPIHPGYSHPGHNSDLM 130
 QY 156 LILNRRIRPTKDVPRINVSCHPSAGTKCLVSGWGTTKSPQVHPKVLQCLINISVLSQK 215
 DB 131 LILNRRIRPTKDVPRINVSCHPSAGTKCLVSGWGTTKSPQVHPKVLQCLINISVLSQK 190
 QY 216 RCEDAYPRQIDDTMFCAGDKAGRSDSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 275
 DB 191 RCEDAYPRQIDDTMFCAGDKAGRSDSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 250
 QY 276 TNLCKFTKWIQETIQANS 293
 DB 251 TNLCKFTKWIQETIQANS 268

RESULT 5
 US-09-764-762-1
 ; Sequence 1, Application US/09764762
 ; Patent No. 6472195
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: NOVEL KALLIKREIN
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive

CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/764,762
 FILING DATE: 16-Jan-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/210,084
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0252 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 268 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: KERANOT02
 CLONE: 820694
 US-09-764-762-1

Query Match 72.9%; Score 1173; DB 4; Length 268;
 Best Local Similarity 85.7%; Pred. No. 1.2e-116;
 Matches 221; Conservative 8; Mismatches 25; Indels 4; Gaps 2;
 QY 38 PSNTVPSGNSQ--DLGAGAGEDARDSDSSRIINGSDCDMHTQPMQAAALLLRPNOLYCGA 95
 DB 13 PLTPCPLGATRTWELGPKTFCGRMTAAASMSMDPTAICT--PSRQQAALLLRPNOLYCGA 70
 QY 96 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIPIHPGYSHPGHNSDLM 155
 DB 71 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIPIHPGYSHPGHNSDLM 130
 QY 156 LILNRRIRPTKDVPRINVSCHPSAGTKCLVSGWGTTKSPQVHPKVLQCLINISVLSQK 215
 DB 131 LILNRRIRPTKDVPRINVSCHPSAGTKCLVSGWGTTKSPQVHPKVLQCLINISVLSQK 190
 QY 216 RCEDAYPRQIDDTMFCAGDKAGRSDSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 275
 DB 191 RCEDAYPRQIDDTMFCAGDKAGRSDSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 250
 QY 276 TNLCKFTKWIQETIQANS 293
 DB 251 TNLCKFTKWIQETIQANS 268

RESULT 6
 US-09-439-313-525
 ; Sequence 525, Application US/09439313
 ; Patent No. 6329505
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Devin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan Louise
 ; APPLICANT: Jiang Yuqi
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Kalos, Michael
 ; APPLICANT: Fanger, Gary

Db 27 --SCSIIINGEDCSPHSQPWAALVME-NELFCSGVLVHPQWVLSAAHCFQNSYTIIGLGL 83
Qy 121 YLSFVYVESGQMFQGVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHPCS 180
Db 84 HSLEADQEPGSMQVEASLSVRHPEYRPLANDMLIKLDESVSSEDTIRSIASQCPT 143
Qy 181 AGTKCLVSGWGTTSKSPQVHPFKVLOCLNLSVLSOKRCEADAYPRQIDDTMFCA-GDKAGRD 239
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLDPLVHPFMSFCAAGGQDQKD 201
Qy 240 SCQDGSQGVVVCNGLSGLVSGWGDYPCARPNRPGVYTNLCKFTKWIOTIOAN 292
Db 202 SCNGDSGGPLICNGYLOGLVSGFGKAPCGQGVGVYTNLCKFTKWIOTIOAN 254

RESULT 9

US-09-679-426-525
; Sequence 525, Application US/09679426
; Patent No. 6759515

GENERAL INFORMATION:

; APPLICANT: Xu Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; CURRENT FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-679-426-525

Query Match 41.7%; Score 671; DB 4; Length 254;
Best Local Similarity 45.7%; Pred. No. 3.4e-63;
Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNVDSCHPSNTVPVSGNQDLGAGAGEDARS 60
Db 1 MATAGNPWGFGLYLI-----LGV-----AGSLVSG----- 26
Qy 61 DSSRIINGSDCDMHTOPWQAALLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Db 27 --SCSIIINGEDCSPHSQPWAALVME-NELFCSGVLVHPQWVLSAAHCFQNSYTIIGLGL 83
Qy 121 YLSFVYVESGQMFQGVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHPCS 180
Db 84 HSLEADQEPGSMQVEASLSVRHPEYRPLANDMLIKLDESVSSEDTIRSIASQCPT 143
Qy 181 AGTKCLVSGWGTTSKSPQVHPFKVLOCLNLSVLSOKRCEADAYPRQIDDTMFCA-GDKAGRD 239
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLDPLVHPFMSFCAAGGQDQKD 201
Qy 240 SCQDGSQGVVVCNGLSGLVSGWGDYPCARPNRPGVYTNLCKFTKWIOTIOAN 292
Db 202 SCNGDSGGPLICNGYLOGLVSGFGKAPCGQGVGVYTNLCKFTKWIOTIOAN 254

RESULT 10

US-09-759-143-525
; Sequence 525, Application US/09759143
; Patent No. 6800746

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-525

Query Match 41.7%; Score 671; DB 4; Length 254;

Best Local Similarity 45.7%; Pred. No. 3.4e-63;

Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNVDSCHPSNTVPVSGNQDLGAGAGEDARS 60
Db 1 MATAGNPWGFGLYLI-----LGV-----AGSLVSG----- 26
Qy 61 DSSRIINGSDCDMHTOPWQAALLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Db 27 --SCSIIINGEDCSPHSQPWAALVME-NELFCSGVLVHPQWVLSAAHCFQNSYTIIGLGL 83
Qy 121 YLSFVYVESGQMFQGVKSIPHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHPCS 180
Db 84 HSLEADQEPGSMQVEASLSVRHPEYRPLANDMLIKLDESVSSEDTIRSIASQCPT 143
Qy 181 AGTKCLVSGWGTTSKSPQVHPFKVLOCLNLSVLSOKRCEADAYPRQIDDTMFCA-GDKAGRD 239
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLDPLVHPFMSFCAAGGQDQKD 201
Qy 240 SCQDGSQGVVVCNGLSGLVSGWGDYPCARPNRPGVYTNLCKFTKWIOTIOAN 292
Db 202 SCNGDSGGPLICNGYLOGLVSGFGKAPCGQGVGVYTNLCKFTKWIOTIOAN 254

RESULT 11

US-09-651-236-525
; Sequence 525, Application US/09651236
; Patent No. 6818751

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.

```
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42718C18
; CURRENT APPLICATION NUMBER: US/09/651,236
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 865
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
;
US-09-651-236-525

Query Match 41.7%; Score 671; DB 4; Length 254;
Best Local Similarity 45.7%; Pred. No. 3.4e-63;
Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;

Qy 1 MATARPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MATAGNPGWFLGYLI-----LGV-----AGSLVSG----- 26

Qy 61 DDSSRIINGSDCDMHTQPMQAAALLRPOLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27 --SCSQIINGEDCSPHSQPMQAAALVME--NELFCSGLVLPQWLVLSATHCFQNSYTIIGLGL 83

Qy 121 YSLSPVYESQQMFQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVRPINVSSHCP 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 HSLEADQEPGSMQVSEASLSVRHPEYNNRPLLDMLIKLDESSESDTIRISIASQCPT 143

Qy 181 AGTKCLVSGWGTTSKSPQVHPKVLQCLNISVLQKRCEDAYPROIDDTMFCA-GDKAGRD 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLYDPLYPHSPMFCAGGGQDQKD 201

Qy 240 SCQDGGGPVVCNGSLQGLVSGWDPYPCARPNNRPGVYTNLCKFTKWIQETIOAN 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 SCNGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254
```

```
RESULT 12
US-09-439-313-523
; Sequence 523, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 523
```

```
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-439-313-523

Query Match 41.4%; Score 666; DB 3; Length 254;
Best Local Similarity 45.4%; Pred. No. 1.2e-62;
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

Qy 1 MATARPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MATAGNPGWFLGYLI-----LGV-----AGSLVSG----- 26

Qy 61 DDSSRIINGSDCDMHTQPMQAAALLRPOLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27 --SCSQIINGEDCSPHSQPMQAAALVME--NELFCSGLVLPQWLVLSATHCFQNSYTIIGLGL 83

Qy 121 YSLSPVYESQQMFQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVRPINVSSHCP 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 HSLEADQEPGSMQVSEASLSVRHPEYNNRPLLDMLIKLDESSESDTIRISIASQCPT 143

Qy 181 AGTKCLVSGWGTTSKSPQVHPKVLQCLNISVLQKRCEDAYPROIDDTMFCA-GDKAGRD 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLYDPLYPHSPMFCAGGGQDQXD 201

Qy 240 SCQDGGGPVVCNGSLQGLVSGWDPYPCARPNNRPGVYTNLCKFTKWIQETIOAN 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 SCNGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254
```

```
RESULT 13
US-09-636-215-523
; Sequence 523, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = any amino acid
```

US-09-636-215-523

Query Match 41.4%; Score 666; DB 4; Length 254;
Best Local Similarity 45.4%; Pred. No. 1.2e-62;
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

```
QY 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTVPVSGNQDLGAGAGEDARS 60
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

QY 61 DDSSRIINGSDCDMHTOPWQAALLLRPNQYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
DB 27 --SCSQIINGDCSPHSQPWQAALVME-NELFCGVLVHPQWLVLSATHCFQNSYITIGLGL 83

QY 121 YLSLPVYESGQMGQGVKSIHPHGVSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPCS 180
DB 84 HSLADQEPGQMGVSEASVSRPEYRNPRLANDLMLIKLDSVSESDTIRISIASOQPT 143

QY 181 AGTKCLVSGWGTTSKPVHFKVQLCLNISVLQKRCEDAYPRQIDDTMFCA-GDKAGRD 239
DB 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVVSEVCSKLYDPLVHPSMFCAGGQXQXD 201

QY 240 SCQDSGPPVNCVCSLQSLVSGWGYPCARPNNRPGVYTNLCKFTKWIQTIQAN 292
DB 202 SCNGDSGGPLICNGYLOGLVSGFKAPCGQGVGVYTNLCKFTWIEKTVQAS 254
```

RESULT 14

US-09-685-166A-523
; Sequence 523, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = 'any amino acid'
US-09-685-166A-523

Query Match 41.4%; Score 666; DB 4; Length 254;
Best Local Similarity 45.4%; Pred. No. 1.2e-62;
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

```
QY 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTVPVSGNQDLGAGAGEDARS 60
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26
```

```
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

QY 61 DDSSRIINGSDCDMHTOPWQAALLLRPNQYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
DB 27 --SCSQIINGDCSPHSQPWQAALVME-NELFCGVLVHPQWLVLSATHCFQNSYITIGLGL 83

QY 121 YLSLPVYESGQMGQGVKSIHPHGVSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPCS 180
DB 84 HSLADQEPGQMGVSEASVSRPEYRNPRLANDLMLIKLDSVSESDTIRISIASOQPT 143

QY 181 AGTKCLVSGWGTTSKPVHFKVQLCLNISVLQKRCEDAYPRQIDDTMFCA-GDKAGRD 239
DB 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVVSEVCSKLYDPLVHPSMFCAGGQXQXD 201

QY 240 SCQDSGPPVNCVCSLQSLVSGWGYPCARPNNRPGVYTNLCKFTKWIQTIQAN 292
DB 202 SCNGDSGGPLICNGYLOGLVSGFKAPCGQGVGVYTNLCKFTWIEKTVQAS 254
```

RESULT 15

US-09-679-426-523
; Sequence 523, Application US/09679426
; Patent No. 6759515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; CURRENT FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = any amino acid
US-09-679-426-523

Query Match 41.4%; Score 666; DB 4; Length 254;
Best Local Similarity 45.4%; Pred. No. 1.2e-62;
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

```
QY 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTVPVSGNQDLGAGAGEDARS 60
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

QY 61 DDSSRIINGSDCDMHTOPWQAALLLRPNQYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
DB 27 --SCSQIINGDCSPHSQPWQAALVME-NELFCGVLVHPQWLVLSATHCFQNSYITIGLGL 83

QY 121 YLSLPVYESGQMGQGVKSIHPHGVSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPCS 180
DB 84 HSLADQEPGQMGVSEASVSRPEYRNPRLANDLMLIKLDSVSESDTIRISIASOQPT 143
```

Db 84 HSLEADQEPGSQMYEASLSVRHPYRNRPLLANDLMLIKLDESVSSESDTIRSIASQCPT 143
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVL SOKRCEDAYPRO1DDTMFCA-GDKAGR D 239
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVVSEVCSKLYDPLYHPSMFCAGGQXQXD 201
Qy 240 SCQGDGSGPVVCGSLQGLVSWGDPYPCARNRPGVYTNLCKFTKWIQETIQAN 292
Db 202 SCNGDSGGPLICNGYLQGLVSFGKAPCGQGVGPGVYTNLCKFTWIEKTVQAS 254

Search completed: July 4, 2005, 15:07:26
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 09:55:49 ; Search time 400 Seconds
(without alignments)
282.595 Million cell updates/sec

Title: US-09-936-271c-14
Perfect score: 1608
Sequence: 1 MATARPPMMVLCALITALL.....VYTNLCRFTKWIQTIQANS 293

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	293	9	US-09-739-907-82
2	1608	100.0	293	9	Sequence 82, Appl
3	1608	100.0	293	9	Sequence 97, Appl
4	1608	100.0	293	9	Sequence 309, Appl
5	1608	100.0	293	9	Sequence 309, Appl
6	1608	100.0	293	9	Sequence 309, Appl
7	1608	100.0	293	9	Sequence 309, Appl
8	1608	100.0	293	9	Sequence 309, Appl
9	1608	100.0	293	9	Sequence 309, Appl
10	1608	100.0	293	9	Sequence 309, Appl
11	1608	100.0	293	9	Sequence 309, Appl

12	1608	100.0	293	9	US-09-993-604-309
13	1608	100.0	293	9	Sequence 309, Appl
14	1608	100.0	293	9	US-09-990-456-309
15	1608	100.0	293	9	Sequence 309, Appl
16	1608	100.0	293	9	US-09-989-721-309
17	1608	100.0	293	9	Sequence 309, Appl
18	1608	100.0	293	9	US-09-992-598-309
19	1608	100.0	293	9	Sequence 309, Appl
20	1608	100.0	293	9	US-09-989-723-309
21	1608	100.0	293	9	Sequence 309, Appl
22	1608	100.0	293	9	US-09-989-723-309
23	1608	100.0	293	9	Sequence 309, Appl
24	1608	100.0	293	10	US-09-989-724-309
25	1608	100.0	293	10	Sequence 309, Appl
26	1608	100.0	293	10	US-09-989-728-309
27	1608	100.0	293	10	Sequence 309, Appl
28	1608	100.0	293	10	US-09-990-441-309
29	1608	100.0	293	10	Sequence 309, Appl
30	1608	100.0	293	10	US-09-997-428-309
31	1608	100.0	293	10	Sequence 309, Appl
32	1608	100.0	293	10	US-09-997-666-309
33	1608	100.0	293	10	Sequence 309, Appl
34	1608	100.0	293	10	US-09-990-438-309
35	1608	100.0	293	10	Sequence 309, Appl
36	1608	100.0	293	10	US-09-990-562-309
37	1608	100.0	293	10	Sequence 309, Appl
38	1608	100.0	293	10	US-09-990-711-309
39	1608	100.0	293	10	Sequence 309, Appl
40	1608	100.0	293	10	US-09-997-573-309
41	1608	100.0	293	10	Sequence 309, Appl
42	1608	100.0	293	10	US-09-991-172-309
43	1608	100.0	293	10	Sequence 309, Appl
44	1608	100.0	293	10	US-09-990-437-309
45	1608	100.0	293	10	Sequence 309, Appl

ALIGNMENTS

RESULT 1
US-09-739-907-82
; Sequence 82, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-907-82

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTVPGSGNODLCAGGEDARS 60
|||||

Db 1 MATARPPMMVLCALITALLIGVTEHVLIANNVSCDHPSTNTVPSGNSQDLGAGAGEDARS 60
Qy 61 DSSSRRIINGSDCDMHTQPWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
Db 61 DSSSRRIINGSDCDMHTQPWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
Qy 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNDLMLIKLNRRIRPTKDVPRPINVSSHCP 180
Db 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNDLMLIKLNRRIRPTKDVPRPINVSSHCP 180
Qy 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKAGRDS 240
Qy 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293
Db 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293

RESULT 2
US-09-739-907-97
; Sequence 97, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 97
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-907-97

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLIANNVSCDHPSTNTVPSGNSQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLIGVTEHVLIANNVSCDHPSTNTVPSGNSQDLGAGAGEDARS 60
Qy 61 DSSSRRIINGSDCDMHTQPWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
Db 61 DSSSRRIINGSDCDMHTQPWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
Qy 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNDLMLIKLNRRIRPTKDVPRPINVSSHCP 180
Db 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNDLMLIKLNRRIRPTKDVPRPINVSSHCP 180
Qy 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKAGRDS 240
Qy 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293
Db 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293

RESULT 3

US-09-989-722-309
; Sequence 309, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C63
; CURRENT APPLICATION NUMBER: US/09/989,722
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355

; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Qy 61 DSSSRRIINGSDCDMHTQPWQAALLLRPNOLYCCAVLVHPQWLLTAACHCKKVFVRVLGH 120
Db 61 DSSSRRIINGSDCDMHTQPWQAALLLRPNOLYCCAVLVHPQWLLTAACHCKKVFVRVLGH 120
Qy 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRIPTKDVPRINVSCHPCS 180
Db 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRIPTKDVPRINVSCHPCS 180
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

Db 181 AGTKCLVSGWGTTSKPVQHPFKVLQCLNISVLQKRCEDAYPRQIDDTMFCAQDKAGRDS 240
Qy 241 CQDSDGGPVVNCVSGLSGLVSGDYPFCARPNRPGVYTNLCKFTKWIQTIQANS 293
Db 241 CQDSDGGPVVNCVSGLSGLVSGDYPFCARPNRPGVYTNLCKFTKWIQTIQANS 293

RESULT 4
US-09-989-723-309
; Sequence 309, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gieritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC82
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19

```

; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPWVLCALITALLGVTEHVLANNVSDCHPSNTVPSGNSQDLGAGAGEDARS 60
DB 1 MATARPPWVLCALITALLGVTEHVLANNVSDCHPSNTVPSGNSQDLGAGAGEDARS 60

; 61 DDSSRIINGSDCDMHTOPWQAALLRPNOLYCGAVLVHPOMLLTAACRKKVFRVLGH 120
; 61 DDSSRIINGSDCDMHTOPWQAALLRPNOLYCGAVLVHPOMLLTAACRKKVFRVLGH 120
; 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNSNDLMLIKLNRRIRPTKDVREINVSCHPS 180
; 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNSNDLMLIKLNRRIRPTKDVREINVSCHPS 180
; 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
; 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
; 241 CQDGGPVCVNGSLQGLVSGMDYPCARNRPGVYTNLCFTKWIQETIOANS 293
; 241 CQDGGPVCVNGSLQGLVSGMDYPCARNRPGVYTNLCFTKWIQETIOANS 293

RESULT 5
US-09-989-279-309
; Sequence 309, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
```

1	PRIOR FILING DATE: 1998-06-18	60/089908
2	PRIOR APPLICATION NUMBER: 60/089908	
3	PRIOR FILING DATE: 1998-06-18	
4	PRIOR APPLICATION NUMBER: 60/089947	
5	PRIOR FILING DATE: 1998-06-19	
6	PRIOR APPLICATION NUMBER: 60/089948	
7	PRIOR FILING DATE: 1998-06-19	
8	PRIOR APPLICATION NUMBER: 60/089952	
9	PRIOR FILING DATE: 1998-06-19	
10	PRIOR APPLICATION NUMBER: 60/090246	
11	PRIOR FILING DATE: 1998-06-22	
12	PRIOR APPLICATION NUMBER: 60/090252	
13	PRIOR FILING DATE: 1998-06-22	
14	PRIOR APPLICATION NUMBER: 60/090254	
15	PRIOR FILING DATE: 1998-06-22	
16	PRIOR APPLICATION NUMBER: 60/090349	
17	PRIOR FILING DATE: 1998-06-23	
18	PRIOR APPLICATION NUMBER: 60/090355	
19	PRIOR FILING DATE: 1998-06-23	
20	PRIOR APPLICATION NUMBER: 60/090429	
21	PRIOR FILING DATE: 1998-06-24	
22	PRIOR APPLICATION NUMBER: 60/090431	
23	PRIOR FILING DATE: 1998-06-24	
24	PRIOR APPLICATION NUMBER: 60/090435	
25	PRIOR FILING DATE: 1998-06-24	
26	PRIOR APPLICATION NUMBER: 60/090444	
27	PRIOR FILING DATE: 1998-06-24	
28	PRIOR APPLICATION NUMBER: 60/090445	
29	PRIOR FILING DATE: 1998-06-24	
30	PRIOR APPLICATION NUMBER: 60/090472	
31	PRIOR FILING DATE: 1998-06-24	
32	PRIOR APPLICATION NUMBER: 60/090535	
33	PRIOR FILING DATE: 1998-06-24	
34	PRIOR APPLICATION NUMBER: 60/090540	
35	PRIOR FILING DATE: 1998-06-24	
36	PRIOR APPLICATION NUMBER: 60/090542	
37	PRIOR FILING DATE: 1998-06-24	
38	PRIOR APPLICATION NUMBER: 60/090557	
39	PRIOR FILING DATE: 1998-06-24	
40	PRIOR APPLICATION NUMBER: 60/090676	
41	PRIOR FILING DATE: 1998-06-25	
42	PRIOR APPLICATION NUMBER: 60/090678	
43	PRIOR FILING DATE: 1998-06-25	
44	PRIOR APPLICATION NUMBER: 60/090690	
45	PRIOR FILING DATE: 1998-06-25	
46	PRIOR APPLICATION NUMBER: 60/090694	
47	PRIOR FILING DATE: 1998-06-25	
48	PRIOR APPLICATION NUMBER: 60/090695	
49	PRIOR FILING DATE: 1998-06-25	
50	PRIOR APPLICATION NUMBER: 60/090696	
51	PRIOR FILING DATE: 1998-06-25	
52	PRIOR APPLICATION NUMBER: 60/090862	
53	PRIOR FILING DATE: 1998-06-25	
54	PRIOR APPLICATION NUMBER: 60/090863	
55	PRIOR FILING DATE: 1998-06-26	
56	PRIOR APPLICATION NUMBER: 60/091360	
57	PRIOR FILING DATE: 1998-07-01	
58	PRIOR APPLICATION NUMBER: 60/091478	
59	PRIOR FILING DATE: 1998-07-02	
60	PRIOR APPLICATION NUMBER: 60/091544	
61	PRIOR FILING DATE: 1998-07-01	
62	PRIOR APPLICATION NUMBER: 60/091519	
63	PRIOR FILING DATE: 1998-07-02	
64	PRIOR APPLICATION NUMBER: 60/091626	
65	PRIOR FILING DATE: 1998-07-02	
66	PRIOR APPLICATION NUMBER: 60/091633	
67	PRIOR FILING DATE: 1998-07-02	
68	PRIOR APPLICATION NUMBER: 60/091978	
69	PRIOR FILING DATE: 1998-07-07	
70	PRIOR APPLICATION NUMBER: 60/091982	
71	PRIOR FILING DATE: 1998-07-07	
72	PRIOR APPLICATION NUMBER: 60/092182	
73	PRIOR FILING DATE: 1998-07-09	

```
Query Match          100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9,1e-150; Indels 0; Gaps 0;
Matches 293; Conservative 0; Mismatches 0;

QY 1 MATARPPWVLCALITALLGVTEHVLANNVSDHPSNTVPVSGNQDLGAGAGEDARS 60
   |||||
Db 1 MATARPPWVLCALITALLGVTEHVLANNVSDHPSNTVPVSGNQDLGAGAGEDARS 60
   |||||

QY 61 DSSSRRIINGSCDHTQWQAALLRPNOLYCGAVLHPQWLLTAACHRKKVFRVLGH 120
   |||||
Db 61 DSSSRRIINGSCDHTQWQAALLRPNOLYCGAVLHPQWLLTAACHRKKVFRVLGH 120
   |||||

QY 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHSHNDMLIKLNRRIPTKDVPRPINVSSHCPs 180
   |||||
Db 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHSHNDMLIKLNRRIPTKDVPRPINVSSHCPs 180
   |||||

QY 181 AGTKCLVSGWGTTPKSPQVHPFKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
   |||||
Db 181 AGTKCLVSGWGTTPKSPQVHPFKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
   |||||

QY 241 CGDSDGGPVVNCVSGSLQGLVSGWGDYFCARPNRPGVVTNLCKFTKWTQETIQANS 293
   |||||
Db 241 CGDSDGGPVVNCVSGSLQGLVSGWGDYFCARPNRPGVVTNLCKFTKWTQETIQANS 293
   |||||
```

RESULT 6

US-09-989-727-309

; Sequence 309, Application US/09989727

; Patent No. US20020072497A1

; GENERAL INFORMATION:

```
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-06-17
```

```
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
```

; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626

; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150; Indels 0; Gaps 0;
Matches 293; Conservative 0; Mismatches 0;

Qy	1	MATARPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPSGSNQDLGAGAGEDARS	60
Db	1	MATARPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPSGSNQDLGAGAGEDARS	60
Qy	61	DDSSRIINGSDCDMHTQPWQAALLLRPNOLYCGAVLVHPQWLLTAHAHCRKKVPRVRLGH	120
Db	61	DDSSRIINGSDCDMHTQPWQAALLLRPNOLYCGAVLVHPQWLLTAHAHCRKKVPRVRLGH	120
Qy	121	YSLSPVYESGOQMPQGVKSTPHPGYSHPGHNDMLIKLNRIRPTKDVPRINVSCHPS	180
Db	121	YSLSPVYESGOQMPQGVKSTPHPGYSHPGHNDMLIKLNRIRPTKDVPRINVSCHPS	180
Qy	181	AGTKCLVSGWGTTKSPQVHPFKVLQCLNLSVLSOKRCEDAYPRQIDDTMFCAGDKAGRDS	240
Db	181	AGTKCLVSGWGTTKSPQVHPFKVLQCLNLSVLSOKRCEDAYPRQIDDTMFCAGDKAGRDS	240
Qy	241	CQDGGGPPVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCCKFTKWIQETIQANS	293
Db	241	CQDGGGPPVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCCKFTKWIQETIQANS	293

RESULT 7

US-09-989-731-309
; Sequence 309, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: KJavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C70
; CURRENT APPLICATION NUMBER: US/09/989,731
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

1	PRIOR APPLICATION NUMBER: 60/065186
2	PRIOR FILING DATE: 1997-11-12
3	PRIOR APPLICATION NUMBER: 60/065311
4	PRIOR FILING DATE: 1997-11-13
5	PRIOR APPLICATION NUMBER: 60/066770
6	PRIOR FILING DATE: 1997-11-24
7	PRIOR APPLICATION NUMBER: 60/075945
8	PRIOR FILING DATE: 1998-02-25
9	PRIOR APPLICATION NUMBER: 60/078910
10	PRIOR FILING DATE: 1998-03-20
11	PRIOR APPLICATION NUMBER: 60/083322
12	PRIOR FILING DATE: 1998-04-28
13	PRIOR APPLICATION NUMBER: 60/084600
14	PRIOR FILING DATE: 1998-05-07
15	PRIOR APPLICATION NUMBER: 60/087106
16	PRIOR FILING DATE: 1998-05-28
17	PRIOR APPLICATION NUMBER: 60/087607
18	PRIOR FILING DATE: 1998-06-02
19	PRIOR APPLICATION NUMBER: 60/087609
20	PRIOR FILING DATE: 1998-06-02
21	PRIOR APPLICATION NUMBER: 60/087759
22	PRIOR FILING DATE: 1998-06-02
23	PRIOR APPLICATION NUMBER: 60/087827
24	PRIOR FILING DATE: 1998-06-03
25	PRIOR APPLICATION NUMBER: 60/088021
26	PRIOR FILING DATE: 1998-06-04
27	PRIOR APPLICATION NUMBER: 60/088025
28	PRIOR FILING DATE: 1998-06-04
29	PRIOR APPLICATION NUMBER: 60/088026
30	PRIOR FILING DATE: 1998-06-04
31	PRIOR APPLICATION NUMBER: 60/088028
32	PRIOR FILING DATE: 1998-06-04
33	PRIOR APPLICATION NUMBER: 60/088029
34	PRIOR FILING DATE: 1998-06-04
35	PRIOR APPLICATION NUMBER: 60/088030
36	PRIOR FILING DATE: 1998-06-04
37	PRIOR APPLICATION NUMBER: 60/088033
38	PRIOR FILING DATE: 1998-06-04
39	PRIOR APPLICATION NUMBER: 60/088326
40	PRIOR FILING DATE: 1998-06-04
41	PRIOR APPLICATION NUMBER: 60/088167
42	PRIOR FILING DATE: 1998-06-05
43	PRIOR APPLICATION NUMBER: 60/088202
44	PRIOR FILING DATE: 1998-06-05
45	PRIOR APPLICATION NUMBER: 60/088212
46	PRIOR FILING DATE: 1998-06-05
47	PRIOR APPLICATION NUMBER: 60/088217
48	PRIOR FILING DATE: 1998-06-05
49	PRIOR APPLICATION NUMBER: 60/088655
50	PRIOR FILING DATE: 1998-06-09
51	PRIOR APPLICATION NUMBER: 60/088734
52	PRIOR FILING DATE: 1998-06-10
53	PRIOR APPLICATION NUMBER: 60/088738
54	PRIOR FILING DATE: 1998-06-10
55	PRIOR APPLICATION NUMBER: 60/088742
56	PRIOR FILING DATE: 1998-06-10
57	PRIOR APPLICATION NUMBER: 60/088826
58	PRIOR FILING DATE: 1998-06-10
59	PRIOR APPLICATION NUMBER: 60/088810
60	PRIOR FILING DATE: 1998-06-10
61	PRIOR APPLICATION NUMBER: 60/088858
62	PRIOR FILING DATE: 1998-06-11
63	PRIOR APPLICATION NUMBER: 60/088861
64	PRIOR FILING DATE: 1998-06-11
65	PRIOR APPLICATION NUMBER: 60/088876
66	PRIOR FILING DATE: 1998-06-11
67	PRIOR APPLICATION NUMBER: 60/089105
68	PRIOR FILING DATE: 1998-06-12
69	PRIOR APPLICATION NUMBER: 60/089440
70	PRIOR FILING DATE: 1998-06-16
71	PRIOR APPLICATION NUMBER: 60/089512

7	PRIOR FILING DATE: 1998-06-16
7	PRIOR APPLICATION NUMBER: 60/089514
7	PRIOR FILING DATE: 1998-06-16
7	PRIOR APPLICATION NUMBER: 60/089532
7	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089538
7	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089598
7	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089599
7	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089600
7	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089653
7	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089801
7	PRIOR FILING DATE: 1998-06-18
7	PRIOR APPLICATION NUMBER: 60/089907
7	PRIOR FILING DATE: 1998-06-18
7	PRIOR APPLICATION NUMBER: 60/089908
7	PRIOR FILING DATE: 1998-06-18
7	PRIOR APPLICATION NUMBER: 60/089947
7	PRIOR FILING DATE: 1998-06-19
7	PRIOR APPLICATION NUMBER: 60/089948
7	PRIOR FILING DATE: 1998-06-19
7	PRIOR APPLICATION NUMBER: 60/089952
7	PRIOR FILING DATE: 1998-06-19
7	PRIOR APPLICATION NUMBER: 60/090246
7	PRIOR FILING DATE: 1998-06-22
7	PRIOR APPLICATION NUMBER: 60/090252
7	PRIOR FILING DATE: 1998-06-22
7	PRIOR APPLICATION NUMBER: 60/090254
7	PRIOR FILING DATE: 1998-06-22
7	PRIOR APPLICATION NUMBER: 60/090349
7	PRIOR FILING DATE: 1998-06-23
7	PRIOR APPLICATION NUMBER: 60/090355
7	PRIOR FILING DATE: 1998-06-23
7	PRIOR APPLICATION NUMBER: 60/090429
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090444
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090445
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090472
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090535
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090540
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090542
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090557
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090676
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090678
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090690
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090694
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090695
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090696
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090862
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/090863
7	PRIOR FILING DATE: 1998-06-26

```
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRITNGSDCMHTQWQAALLRPNQLYCGAVLVHPQWLLTAHCRKKVPRVLGH 120
Db 61 DSSSRITNGSDCMHTQWQAALLRPNQLYCGAVLVHPQWLLTAHCRKKVPRVLGH 120

Qy 121 YLSPPVYESGQMGQGVKSIHPHGVSHPGHNDMLKLNRRIRPTKDVPRINVSCHPS 180
Db 121 YLSPPVYESGQMGQGVKSIHPHGVSHPGHNDMLKLNRRIRPTKDVPRINVSCHPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVLQCLNISVLQKRCEDAYPRIDDTMFCAQDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFVKVLQCLNISVLQKRCEDAYPRIDDTMFCAQDKAGRDS 240

Qy 241 CQDGGGPPVVCNGLQGLVSGNDYPCARPNGPVYTNLCCKTKWIQETIQANS 293
Db 241 CQDGGGPPVVCNGLQGLVSGNDYPCARPNGPVYTNLCCKTKWIQETIQANS 293

RESULT 8
US-09-989-732-309
; Sequence 309, Application US/09989732
; Patent No. US20020123463A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
```

```
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC57
; CURRENT APPLICATION NUMBER: US/09/989, 732
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
```


; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9, 1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLLGVTGTEHVLANNDVSCDHPSTNTPVSGSNODLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLLGVTGTEHVLANNDVSCDHPSTNTPVSGSNODLGAGAGEDARS 60

Qy 61 DSSSRRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120
Db 61 DSSSRRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120

Qy 121 YLSLPVYESGQMFQGVKSIHPGYSHPGHSHNDLMLIKLNRRIPTKDVRIINVSCHCPS 180
Db 121 YLSLPVYESGQMFQGVKSIHPGYSHPGHSHNDLMLIKLNRRIPTKDVRIINVSCHCPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

Qy 241 CQDGGGPPVVCNCSLQGLVSWGDPYPCARPNGVYTNLCKFTKWIQETIQANS 293
Db 241 CQDGGGPPVVCNCSLQGLVSWGDPYPCARPNGVYTNLCKFTKWIQETIQANS 293

RESULT 9
US-09-931-073-309
; Sequence 309, Application US/09991073
; Patent No. US20020127576A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.

APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secrated and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC15
CURRENT APPLICATION NUMBER: US/09/991.073
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPWVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPWVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRINSGDCMHTQWQAALLRNQLYCGAVLVHPQWLLTAAHCRKKVFRVLGH 120
Db 61 DSSSRINSGDCMHTQWQAALLRNQLYCGAVLVHPQWLLTAAHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGOOMFGQVKSIHPGYSHPGSHNDLMLIKLNRIPTKDVRIINVSCHPS 180
Db 121 YSLSPVYESGOOMFGQVKSIHPGYSHPGSHNDLMLIKLNRIPTKDVRIINVSCHPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPROIDDTMFCAKGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPROIDDTMFCAKGRDS 240

Qy 241 CQDGGGPPVWCVNGSLQGLVSGDYPFCARPNRPVYTNLCCKTKVIQETIQANS 293
Db 241 CQDGGGPPVWCVNGSLQGLVSGDYPFCARPNRPVYTNLCCKTKVIQETIQANS 293

RESULT 10
US-09-936-442-309
Sequence 309, Application US/09990442
Patent No. US20020132252A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C8
CURRENT APPLICATION NUMBER: US/09/990,442
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;

Best Local Similarity 100.0%; Pred. No. 9.1e-150; Mismatches 0; Indels 0; Gaps 0;
Matches 293; Conservative 0;

Qy	1	MATARPMMVLCALITALLGVTEHVLANNVSCDHPSNTVPSGSNODLGAGEDARS	60
Db	1	MATARPMMVLCALITALLGVTEHVLANNVSCDHPSNTVPSGSNODLGAGEDARS	60
Qy	61	DDSSRIINGSDCDMHTQPMQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH	120
Db	61	DDSSRIINGSDCDMHTQPMQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH	120
Qy	121	YSLSPVYESGQMPQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVPRINVSCHPS	180
Db	121	YSLSPVYESGQMPQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVPRINVSCHPS	180
Qy	181	AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEDAYPRIDDTMFCAGDKGRDS	240
Db	181	AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEDAYPRIDDTMFCAGDKGRDS	240
Qy	241	CQGDGGPVVVCNGLQGLVSWGDYPCARPNNRPVYTNLCCKFTKWIQETIQANS	293
Db	241	CQGDGGPVVVCNGLQGLVSWGDYPCARPNNRPVYTNLCCKFTKWIQETIQANS	293

RESULT 11
US-09-991-163-309

; Sequence 309, Application US/09991163
; Patent No. US20020132253A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC17
; CURRENT APPLICATION NUMBER: US/09/991,163
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23

```

; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRIINGSDCDMHTOPWQAALLRRNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120
Db 61 DSSSRIINGSDCDMHTOPWQAALLRRNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMFQGVKSIHPHGVSHPHGNSNDLMLIKLNNRIPTKDVRFINVSSHCPs 180
Db 121 YSLSPVYESGQMFQGVKSIHPHGVSHPHGNSNDLMLIKLNNRIPTKDVRFINVSSHCPs 180

Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQSKRCEDAYPRQIDTMTFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQSKRCEDAYPRQIDTMTFCAGDKAGRDS 240

; Sequence 309, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04

RESULT 12
US-09-993-604-309
; Sequence 309, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
```

; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246

; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;

Best Local Similarity 100.0%; Pred. No. 9.1e-150;

Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Db 1 MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DDSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120

Db 61 DSSSRINGSDCMHTOPWQAALLRNLNQLYCCAVLHPQWLLTAHCRKKVFRVLGH 120
Qy 121 YSLSPVYESGQMFQGVKSIHPHGYSHPHGSHNDMLKLNRRIRPTKDVREINVSCHPS 180
Db 121 YSLSPVYESGQMFQGVKSIHPHGYSHPHGSHNDMLKLNRRIRPTKDVREINVSCHPS 180
Qy 181 AGTKLVSGWTTKSPQVHFVKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240
Db 181 AGTKLVSGWTTKSPQVHFVKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240
Qy 241 CQDSGGPVWNCISGLVSGWDPFCARPNRPVYTNLCKTKWIOETIQANS 293
Db 241 CQDSGGPVWNCISGLVSGWDPFCARPNRPVYTNLCKTKWIOETIQANS 293

RESULT 13
US-09-930-456-309
; Sequence 309, Application US/09990456
; Patent No. US20020137890A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730FIC22
; CURRENT APPLICATION NUMBER: US/09/990,456
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18

;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTVPSGNSQDLGAGAGEDARS 60
DB 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSCDHPSTNTVPSGNSQDLGAGAGEDARS 60

QY 61 DSSSRRIINGSDCDMHTOPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
DB 61 DSSSRRIINGSDCDMHTOPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120

QY 121 YSLSPVYESGQOMFQGVKSIHPGYSHPGHNDLMLKLNRRIRPTKDVRIINVSCHPCS 180
DB 121 YSLSPVYESGQOMFQGVKSIHPGYSHPGHNDLMLKLNRRIRPTKDVRIINVSCHPCS 180

QY 181 AGTKCLVSGMGTTKSPQVHPFKVLQCLNLSVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGMGTTKSPQVHPFKVLQCLNLSVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

QY 241 CQDGGGPPVVCNGSLQGLVSWGDPYPCARPNRPVGYTNLCKTKWIOETIQANS 293
DB 241 CQDGGGPPVVCNGSLQGLVSWGDPYPCARPNRPVGYTNLCKTKWIOETIQANS 293

RESULT 14

US-09-989-721-309
; Sequence 309, Application US/09989721
; Patent No. US20020142961A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C55
; CURRENT APPLICATION NUMBER: US/09/989,721
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20

, PRIOR APPLICATION NUMBER: 60/083322
 , PRIOR FILING DATE: 1998-04-28
 , PRIOR APPLICATION NUMBER: 60/084600
 , PRIOR FILING DATE: 1998-05-07
 , PRIOR APPLICATION NUMBER: 60/087106
 , PRIOR FILING DATE: 1998-05-28
 , PRIOR APPLICATION NUMBER: 60/087607
 , PRIOR FILING DATE: 1998-06-02
 , PRIOR APPLICATION NUMBER: 60/087609
 , PRIOR FILING DATE: 1998-06-02
 , PRIOR APPLICATION NUMBER: 60/087759
 , PRIOR FILING DATE: 1998-06-02
 , PRIOR APPLICATION NUMBER: 60/087827
 , PRIOR FILING DATE: 1998-06-03
 , PRIOR APPLICATION NUMBER: 60/088021
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088025
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088026
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088028
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088029
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088030
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088033
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088326
 , PRIOR FILING DATE: 1998-06-04
 , PRIOR APPLICATION NUMBER: 60/088167
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088202
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088212
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088217
 , PRIOR FILING DATE: 1998-06-05
 , PRIOR APPLICATION NUMBER: 60/088655
 , PRIOR FILING DATE: 1998-06-09
 , PRIOR APPLICATION NUMBER: 60/088734
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088738
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088742
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088810
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088824
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088826
 , PRIOR FILING DATE: 1998-06-10
 , PRIOR APPLICATION NUMBER: 60/088858
 , PRIOR FILING DATE: 1998-06-11
 , PRIOR APPLICATION NUMBER: 60/088861
 , PRIOR FILING DATE: 1998-06-11
 , PRIOR APPLICATION NUMBER: 60/088876
 , PRIOR FILING DATE: 1998-06-11
 , PRIOR APPLICATION NUMBER: 60/0889105
 , PRIOR FILING DATE: 1998-06-12
 , PRIOR APPLICATION NUMBER: 60/089440
 , PRIOR FILING DATE: 1998-06-16
 , PRIOR APPLICATION NUMBER: 60/089512
 , PRIOR FILING DATE: 1998-06-16
 , PRIOR APPLICATION NUMBER: 60/089514
 , PRIOR FILING DATE: 1998-06-16
 , PRIOR APPLICATION NUMBER: 60/089532
 , PRIOR FILING DATE: 1998-06-17
 , PRIOR APPLICATION NUMBER: 60/089538
 , PRIOR FILING DATE: 1998-06-17
 , PRIOR APPLICATION NUMBER: 60/089598
 , PRIOR FILING DATE: 1998-06-17
 , PRIOR APPLICATION NUMBER: 60/089599

```
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMWVLCALITALLIGVTEHVLANDVSCDHESNTVPSSNODLGAGAGEDARS 60
Db 1 MATARPPMMWVLCALITALLIGVTEHVLANDVSCDHESNTVPSSNODLGAGAGEDARS 60
Qy 61 DSSSRRIINGSDCDMHTOPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH 120
Db 61 DSSSRRIINGSDCDMHTOPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH 120
Qy 121 YSLSPVYESGQMGFGVKSIPHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSSHCPs 180
Db 121 YSLSPVYESGQMGFGVKSIPHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSSHCPs 180
Qy 181 AGTKCLVSGWGTTSKSPQVHPFKVLOCINISVLQKRCEDAYPRQIDDDTFMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTSKSPQVHPFKVLOCINISVLQKRCEDAYPRQIDDDTFMFCAGDKAGRDS 240
Qy 241 CQSDSGGPVVCNGSLQGLVSGWDYPCARNRPGVVTNLCKFTKWIQETIQANS 293
Db 241 CQSDSGGPVVCNGSLQGLVSGWDYPCARNRPGVVTNLCKFTKWIQETIQANS 293

RESULT 15
US-09-992-598-309
; Sequence 309, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
```

```
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
```

;
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360

;
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLGVTEHVLANNDVSCDHPSNTVPSSGNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLGVTEHVLANNDVSCDHPSNTVPSSGNQDLGAGAGEDARS 60

Qy 61 DDSSRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120
Db 61 DDSSRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120

Qy 121 YSLSPVYESQQMFQGVKSIHPFGYSHPGHNSNDMLIKLNRRIPTKDVRLPINVSSHCP 180
Db 121 YSLSPVYESQQMFQGVKSIHPFGYSHPGHNSNDMLIKLNRRIPTKDVRLPINVSSHCP 180

Qy 181 AGTKCLVSGWGTTKSPQVHPKVLQCLNLSVLSQKRCEDAYPROIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHPKVLQCLNLSVLSQKRCEDAYPROIDDTMFCAGDKAGRDS 240

Qy 241 CQGDSSGPPVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCCKFTKWIQETIQANS 293
Db 241 CQGDSSGPPVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCCKFTKWIQETIQANS 293

Search completed: July 4, 2005, 15:19:50
Job time : 402 secs

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	627	39.0	260	2	I56559
2	612.5	38.1	233	2	A53968
3	586	36.4	246	1	TRRT1
4	582.5	36.2	247	2	A27547
5	578	35.9	247	1	TRDG
6	571.5	35.5	229	1	TRBOTR
7	569.5	35.4	248	2	S55066
8	563.5	35.0	247	1	A25852
9	562.5	35.0	246	1	TRRT2
10	561.5	34.9	231	1	TRPQTR
11	560	34.8	247	2	S13813
12	558.5	34.7	238	2	S31779
13	555.5	34.5	246	1	TRDGC
14	550	34.2	244	2	A44284
15	544	33.8	246	2	QJ1472
16	538	33.5	261	2	A31136
17	537.5	33.4	246	2	B25528
18	536	33.3	259	2	B31136
19	535	33.3	246	2	QJ1471
20	532.5	33.1	243	2	A35871
21	532.5	33.1	248	2	S55067
22	530	33.0	261	1	A32297
23	529	32.9	261	2	A34079
24	526	32.7	265	1	KOR7P
25	525.5	32.7	247	1	B25852
26	525	32.6	261	2	A29586
27	524.5	32.6	304	2	S33496
28	523.5	32.6	242	2	S49489
29	523.5	32.6	247	2	S05494
30	523.5	32.6	247	2	S05494
31	523.5	32.6	247	2	S05494
32	523.5	32.6	247	2	S05494
33	523.5	32.6	247	2	S05494
34	523.5	32.6	247	2	S05494
35	523.5	32.6	247	2	S05494
36	523.5	32.6	247	2	S05494
37	523.5	32.6	247	2	S05494
38	523.5	32.6	247	2	S05494
39	523.5	32.6	247	2	S05494
40	523.5	32.6	247	2	S05494
41	523.5	32.6	247	2	S05494
42	523.5	32.6	247	2	S05494
43	523.5	32.6	247	2	S05494
44	523.5	32.6	247	2	S05494
45	523.5	32.6	247	2	S05494
46	523.5	32.6	247	2	S05494
47	523.5	32.6	247	2	S05494
48	523.5	32.6	247	2	S05494
49	523.5	32.6	247	2	S05494
50	523.5	32.6	247	2	S05494
51	523.5	32.6	247	2	S05494
52	523.5	32.6	247	2	S05494
53	523.5	32.6	247	2	S05494
54	523.5	32.6	247	2	S05494
55	523.5	32.6	247	2	S05494
56	523.5	32.6	247	2	S05494
57	523.5	32.6	247	2	S05494
58	523.5	32.6	247	2	S05494
59	523.5	32.6	247	2	S05494
60	523.5	32.6	247	2	S05494
61	523.5	32.6	247	2	S05494
62	523.5	32.6	247	2	S05494
63	523.5	32.6	247	2	S05494
64	523.5	32.6	247	2	S05494
65	523.5	32.6	247	2	S05494
66	523.5	32.6	247	2	S05494
67	523.5	32.6	247	2	S05494
68	523.5	32.6	247	2	S05494
69	523.5	32.6	247	2	S05494
70	523.5	32.6	247	2	S05494
71	523.5	32.6	247	2	S05494
72	523.5	32.6	247	2	S05494
73	523.5	32.6	247	2	S05494
74	523.5	32.6	24		

J. Biol. Chem. 269, 19420-19426, 1994

C;Species: Canis lupus familiaris (dog)
 C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
 C;Accession: A26273
 R;Pinsky, S.D.; LaForge, K.S.; Scheele, G.
 Mol. Cell. Biol. 5, 2669-2676, 1985
 A;Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA sequence
 A;Reference number: A26273; MUID:86284628; PMID:3841794
 A;Accession: A26273
 A;Molecule type: mRNA
 A;Residues: 1-247 <PIN>
 A;Cross-references: UNIPROT:P06872; GB:M1589; NID:g164094; PIDN:AAA30899.1; PID:g164095
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F;1-15/Domain: signal sequence #status predicted <SIG>
 F;16-23/Domain: activation peptide #status predicted <APT>
 F;24-247/Product: trypsin, anionic #status predicted <TRY>
 F;24-239/Domain: trypsin homology <TRY>
 F;30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
 F;63,107,200/Active site: His, Asp, Ser #status predicted
 F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 35.9%; Score 578; DB 1; Length 247;
 Best Local Similarity 46.1%; Pred. No. 9.5e-44;
 Matches 113; Conservative 40; Mismatches 84; Indels 8; Gaps 5;

Qy 50 LGAGGEDARDDSSRIINGSDCDMHTPQWQAALLLRPNOLYCGAVLVHPQWLLTAHC 109
 Db 10 LGAAVATPTDDDD---KIVGGYTCEENSVPYQVS--LNAGYHFCGSLSDQWVVAHC 64
 Qy 110 RKVFRVRLGHYLSPVYVESQGMFGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKV 169
 Db 65 YKSIQVRLGEYND-VLEGNEQFINSAKVRHFNYSWILDNDMLIKLSSPAVLNARV 123
 Qy 170 RPNVSHCHPSAGTKCLVSGMTTKSPQVHPKVLQCLINISLVSKRCEDAYPRQIDDTM 229
 Db 124 ATISLPACAPGPTQCLISGNGTSLSSGTNYPELLQCLDAPILTAQCEASYPQGITENM 183
 Qy 230 FCAGD-KAGRDSGCGSGGPPVVCNGSLQGLVSGDYPCARPNRPGVYTNLCFKTKWQET 288
 Db 184 ICAGFLEGGKDCQSGSGPPVVCNGELQGLVSWG-YGCAQKNKPGVYTKVNCVFDWIQST 242
 Qy 289 IQANS 293
 Db 243 IAANS 247

RESULT 6
 TRBOTR
 trypsin (EC 3.4.21.4) precursor - bovine
 N;Contains: trypsinogen
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change 18-Jul-1997
 C;Accession: A90164; A00946; S08774
 R;Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
 Biochem. Biophys. Res. Commun. 24, 346-352, 1966
 A;Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.
 A;Reference number: A90164; MUID:67169848; PMID:5967094
 A;Accession: A90164
 A;Molecule type: protein
 A;Residues: 1-57, 'Q', 59-67, 'Q', 69-150, 'N', 152-176, 'N', 178-229 <MIK>
 R;Hartley, B.S.
 Philos. Trans. R. Soc. Lond. B257, 77-87, 1970
 A;Reference number: A93755
 C;Contents: annotation; revisions
 R;Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
 Biochemistry 14, 1358-1366, 1975
 A;Title: Amino acid sequence of dogfish trypsin.
 A;Reference number: A00950; MUID:75146445; PMID:1092332
 A;Contents: annotation; revisions
 A;Note: the sequence agrees with that shown
 R;Bode, W.; Schwager, P.
 J. Mol. Biol. 98, 693-717, 1975
 A;Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution

A;Reference number: A92954; MUID:76072097; PMID:512
 A;Contents: annotation; x-ray crystallography; binding sites for calcium, substrate, and
 C;Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.
 C;Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a termi
 s pseudotrypsin. A cleavage may also occur after Arg-105.
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F;1-229/Product: trypsinogen #status experimental <ZYM>
 F;1-6/Domain: activation peptide #status experimental <APT>
 F;7-222/Domain: trypsin homology <TRY>
 F;7-132-229/Product: alpha-trypsin #status experimental <MPT>
 F;6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental
 F;13-143,31-47,115-216,122-189,154-166,179-203/Disulfide bonds: #status experimental
 F;46,90,183/Active site: His, Asp, Ser #status experimental
 F;58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental
 F;131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match 35.5%; Score 571.5; DB 1; Length 229;
 Best Local Similarity 46.6%; Pred. No. 3.3e-43;
 Matches 108; Conservative 39; Mismatches 80; Indels 5; Gaps 4;

Qy 62 DSSRIINGSDCDMHTPQWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGHY 121
 Db 2 DDDDKIVGGVTCGANTVPYQVS--LNSGYHFCGSLNSQWVVAACHYKSGIQVRLGED 59
 Qy 122 SLSPVYVESQGMFGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKVDPNPVNSHCHPSA 181
 Db 60 NIN-VVSGNEQFTSASKSIHVHPSYNSNTLNNDMLIKLSAASLSNRVASISLPTSCASA 118
 Qy 182 GTKCLVSGWGTTPQVHPKVLQCLINISLVSKRCEDAYPRQIDDTMFCAGD-KAGRDS 240
 Db 119 GTQCLISGNGTSSGTSTYFDVLKAPILTAQCEASYPQGITSNMFCAGYLEGGKDS 178
 Qy 241 CQDGGGPPVVCNGSLQGLVSGDYPCARPNRPGVYTNLCFKTKWQETIQAN 292
 Db 179 CQDGGGPPVVCNGKLGIVSWGSG-CAQKNKPGVYTKVNCVSVIKQTIASN 229

RESULT 7
 S55066
 trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
 N;Alternate names: trypsinogen II
 C;Species: Gallus gallus (chicken)
 C;Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
 C;Accession: S55066; S72347
 R;Wang, K.; Gan, L.; Lee, I.; Hood, L.
 Biochem. J. 307, 471-479, 1995
 A;Title: Isolation and characterization of the chicken trypsinogen gene family.
 A;Reference number: S55066; MUID:95251611; PMID:7733885
 A;Accession: S55066
 A;Molecule type: mRNA
 A;Residues: 1-248 <WAN1>
 A;Cross-references: UNIPROT:Q90629; EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g6039
 A;Experimental source: clone 2-P29
 A;Accession: S72347
 A;Molecule type: DNA
 A;Residues: 1-248 <WAN2>
 A;Cross-references: EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g603907
 A;Experimental source: clone 2-P29
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F;1-16/Domain: signal sequence #status predicted <SIG>
 F;17-25/Domain: activation peptide #status predicted <APT>
 F;26-248/Product: trypsin II #status predicted <MAT>
 F;26-241/Domain: trypsin homology <TRY>
 F;65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 35.4%; Score 569.5; DB 2; Length 248;
 Best Local Similarity 46.1%; Pred. No. 5.4e-43;
 Matches 112; Conservative 38; Mismatches 86; Indels 7; Gaps 5;

Qy 50 LGAGGEDARDDSSRIINGSDCDMHTPQWQAALLLRPNOLYCGAVLVHPQWLLTAHC 109
 Db 10 LGAAVATPTDDDD---KIVGGYTCEENSVPYQVS--LNAGYHFCGSLSDQWVVAHC 64

```

Db 11 LGAAVAFPGGADD--DKIVGYTTPCHSVFPYQVS--LNSGYHFCGSLNSQWVLSAHC 66
Qy 110 RKVFRVRLGHYSLSPVYESGQOMFQGVKSPHPGYSHPGHSNDMLMLKLNRRIRPTKDV 169
Db 67 YKSIQVRLGHEYNID-VOEDSEVVRSSSVIIRHPKYSSITLNDIMLIKLSA VEYSADI 125
Qy 170 RPIINVSHCHPSAGTKCLVSGWTTKSPQVHPKVLQCLNLSVLSQKRCEDAYPRQIDDTM 229
Db 126 QPIALPSSCAGAKTECLISGWNTLSNGYNYPPELLQCLNAPILSDQCEQAYPGDITSNM 185
Qy 230 FCAGD-KAGRDSGCGSGPVVCGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQET 288
Db 186 ICVGFLEGGKDCQCGSGSPVVCNGELQGI VSWG-IGCALKYPGVYTKVCNVYVDWIQET 244
Qy 289 IQA 291
Db 245 IAA 247

RESULT 8
A:Species: Homo sapiens (man)
N: trypsin (EC 3.4.21.4) I precursor [validated] - human
C: trypsin (EC 3.4.21.4) I precursor - human
C: Species: Homo sapiens (man)
C: Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C: Accession: A25852; B61066; A43988
R: Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Mori, T.; Matsubara, K.
Gene 41, 305-310, 1986
A: Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human
A: Reference number: A91544; MUID: 86221712; PMID: 3011602
A: Accession: A25852
A: Molecule type: mRNA
A: Residues: 1-247 <EM>
A: Cross-references: UNIPROT:P07477; GB:M26212; NID:g521215; PIDN:AAA61231.1; PID:g521216
R: Kilmann, M.; Russick, C.; Marks, W.H.; Borgstroem, A.
Clin. Chim. Acta 184, 31-46, 1989
A: Title: Immunoreactive anionic and cationic trypsin in human serum.
A: Reference number: A61066; MUID: 90091010; PMID: 2598466
A: Accession: B61066
A: Molecule type: protein
A: Residues: 16-43 <KM>
R: Koivunen, E.; Huhtala, M.L.; Stenman, U.H.
J. Biol. Chem. 264, 14095-14099, 1989
A: Title: Human ovarian tumor-associated trypsin. Its purification and characterization f
A: Reference number: A43988; MUID: 89340515; PMID: 2503510
A: Accession: A43988
A: Molecule type: protein
A: Residues: 16-54 <KOI>
A: Experimental source: mucinous ovarian tumor cyst fluid
C: Genetics:
A: Gene: GDB:PRSS1; TRV1
A: Cross-references: GDB:119620; OMIM:276000
A: Map position: 7q35-7q35
A: Note: The human genome contains at least ten trypsin genes or pseudogenes, at least tw
C: Superfamily: trypsin; trypsin homology
C: Keywords: hydrolase; pancreas; phosphoprotein; protein digestion; serine proteinase; z
F: 1-15/Domain: signal sequence #status predicted <SIG>
F: 16-246/Product: trypsin I #status experimental <ZYM>
F: 16-246/Product: trypsin I #status predicted <ENZ>
F: 24-239/Domain: trypsin homology <TRY>
F: 30-160, 48-64, 139-206, 171-185, 196-220/Disulfide bonds: #status predicted
F: 63, 107, 200/Active site: His, Asp, Ser #status predicted
F: 75, 77, 80, 85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 35.0%; Score 563.5; DB 1; Length 247;
Best Local Similarity 46.4%; Pred. No. 1.8e-42;
Matches 108; Conservative 40; Mismatches 80; Indels 5; Gaps 4;

Qy 62 DSSSRINGSDCDMHTQWQALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGHY 121
Db 19 DDDDKIVGGYCNCEENSVFPYQVS--LNSGYHFCGSLNSQWVLSAHCYKSIQVRLGHEH 76

RESULT 10
TRPGTR
trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
N: Contains: trypsinogen

```

```

Qy 122 SLSPVYESGQOMFQGVKSPHPGYSHPGHSNDMLMLKLNRRIRPTKDV RPIINVSHCHPSA 181
Db 77 NIE-VLEGNQFNAAKIIRHPQVDRKTLNDIMLIKLSRAVINARVSTISLTAPPAT 135
Qy 182 GTKCLVSGWTTKSPQVHPKVLQCLNLSVLSQKRCEDAYPRQIDDTMFCAGD-KAGRDS 240
Db 136 GTKCLISGWNTASSGADYDPDELQCLDAPVLSQAKCEASYPGKITSNMFCVGFLEGGKDS 195
Qy 241 CQDSDGSPVVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIOANS 293
Db 196 CQDSDGSPVVCNGSLQGLVSWGD-GCAQKNKPGVYTKVYVYVVKWIKTIAANS 247

RESULT 9
TRRT2
trypsin (EC 3.4.21.4) II precursor - rat
N: Alternate names: trypsinogen II
C: Species: Rattus norvegicus (Norway rat)
C: Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C: Accession: A22657; A00949
R: Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14255-14264, 1984
A: Title: Structure of two related rat pancreatic trypsin genes.
A: Reference number: A22657; MUID: 85054880; PMID: 6094547
A: Accession: A22657
A: Molecule type: DNA
A: Residues: 1-246 <CRA>
A: Cross-references: UNIPROT:P00763
R: MacDonald, R.J.; Stary, S.J.; Swift, G.H.
J. Biol. Chem. 257, 9724-9732, 1982
A: Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
A: Reference number: A00948; MUID: 82265624; PMID: 6896710
A: Accession: A00949
A: Molecule type: mRNA
A: Residues: 9-246 <MAC>
C: Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I mR
C: Genetics:
A: Introns: 14/1; 67/2
A: Superfamily: trypsin; trypsin homology
C: Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F: 1-15/Domain: signal sequence #status predicted <SIG>
F: 16-23/Domain: activation peptide #status predicted <APT>
F: 24-246/Product: trypsin II #status predicted <ENZ>
F: 24-239/Domain: trypsin homology <TRY>
F: 30-160, 48-64, 132-233, 139-206, 171-185/Disulfide bonds: #status predicted
F: 63, 107, 200/Active site: His, Asp, Ser #status predicted
F: 75, 77, 80, 85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 35.0%; Score 562.5; DB 1; Length 246;
Best Local Similarity 45.7%; Pred. No. 2.2e-42;
Matches 106; Conservative 41; Mismatches 80; Indels 5; Gaps 4;

Qy 62 DSSSRINGSDCDMHTQWQALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGHY 121
Db 19 DDDDKIVGGYCNCEENSVFPYQVS--LNSGYHFCGSLNSQWVLSAHCYKSIQVRLGHEH 76
Qy 122 SLSPVYESGQOMFQGVKSPHPGYSHPGHSNDMLMLKLNRRIRPTKDV RPIINVSHCHPSA 181
Db 77 NIN-VLEGNQFNAAKIIRHPQVDRKTLNDIMLIKLSRAVINARVSTISLTAPPAT 135
Qy 182 GTKCLVSGWTTKSPQVHPKVLQCLNLSVLSQKRCEDAYPRQIDDTMFCAGD-KAGRDS 240
Db 136 GTQCLISGWNTLSGNGVNEPDLQLDAPLQADCEASYPGKITSNMFCVGFLEGGKDS 195
Qy 241 CQDSDGSPVVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIOAN 292
Db 196 CQDSDGSPVVCNGSLQGLVSWG-YCCALPDNPGVYTKVYVYVVKWIKTIAAN 246

RESULT 10
TRPGTR
trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
N: Contains: trypsinogen

```


C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 09-Jul-2004
C;Accession: A90641; A90368; A90947
R;Charles, M.; Roversy, M.; Guidoni, A.; Desnuelle, P.
Biochim. Biophys. Acta 69, 115-129, 1963
A;Title: Su le trypsinogène et la trypsine de porc.
A;Reference number: A90641
A;Accession: A90641
A;Molecule type: Protein
A;Residues: 1-10 <CHA>
A;Cross-references: UNIPROT:P00761
R;Hermanson, M.A.; Erickson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 12, 3146-3153, 1973
A;Title: Determination of the amino acid sequence of porcine trypsin by sequenator analysis
A;Reference number: A90368; MUID:73258692; PMID:4738933
A;Accession: A90368
A;Molecule type: Protein
A;Residues: 9-231 <HER>
A;Note: at position 20, Ile and Val occur alternatively
A;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zym
F;1-231/Product: trypsinogen #status experimental <ZYM>
F;1-8/Domain: activation peptide #status experimental <APT>
F;9-231/Product: trypsin #status experimental <MAT>
F;9-224/Domain: trypsin homology <TRY>
F;15-145,33-49,117-218,124-191,156-170,181-205/disulfide bonds: #status predicted
F;48,92,185/Active site: His, Asp, Ser #status predicted
F;60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 34.9%; Score 561.5; DB 1; Length 231;
Best Local Similarity 45.3%; Pred. No. 2.5e-42;
Matches 105; Conservative 39; Mismatches 83; Indels 5; Gaps 4;

Qy 62 DSSSRRIINGSCDMDHPTQWQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH 121
Db 4 DDDDKIVGGYTCANSTIPYQVS--LNSGSHFCGSLNSQWVSAAHCKYSRIQVRLGEH 61

Qy 122 SLSPVYSGQOMFQGVKSIPIHPGVSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPCA 181
Db 62 NID-VLEGNQFVINAAKIITHPNFNGNTLDNDMLIKLSSPATLNSRVATVSLPRSCAA 120

Qy 182 GTKLVSQWGTGKSPQVHPFKVQLCLNISVLQKRCEDAYPRQIDDTMFCAGD-KAGRDS 240
Db 121 GTECLISGNGTKSGSGSPYSLLOCLKAPVLSDSCKSSYPGQITGNMIVGFLGGKDS 180

Qy 241 CQGDGSGPVVCGSLQGLVSGWDYPCARPNNRPGVYTNLCKFTKWIQETIOAN 292
Db 181 CQGDGSGPVVCGSLQGLVSGW-YGCAQKNKPGVYTKVCNVVNIQQTIAAN 231

RESULT 11
S13813
trypsin (EC 3.4.21.4) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S13813
R;Le Huerou, I.; Wicker, C.; Guilloteau, P.; Toullec, R.; Puigserver, A.
Eur. J. Biochem. 193, 767-773, 1990
A;Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic
A;Reference number: S13813; MUID:91065383; PMID:1701147
A;Accession: S13813
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-247 <HUE>
A;Cross-references: UNIPROT:Q29463; EMBL:X54703; NID:g829; PIDN:CAA38513.1; PID:g830
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; protein digestion; serine proteinase
F;24-239/Domain: trypsin homology <TRY>
F;63,107,200/Active site: His, Asp, Ser #status predicted

Query Match 34.8%; Score 560; DB 2; Length 247;
Best Local Similarity 47.1%; Pred. No. 3.7e-42;
Matches 114; Conservative 37; Mismatches 85; Indels 6; Gaps 5;

Qy 53 GAGEDARDSSSRRIINGSCDMDHPTQWQAALLLRPNQLYCGAVLVHPQWLLTAHCRKK 112
Db 11 GAAVAFPSDD-DDKIVGGYTCANSTIPYQVS--LNSGSHFCGSLNSQWVSAAHCKYQY 67

Qy 113 VFRVRLGHYSLSPVYSGQOMFQGVKSIPIHPGVSHPGHNSDMLIKLNRRIRPTKDVPR 172
Db 68 HIQVRLGEYNID-VLEGGEQPIDASKIIRHPKYSSTWLDNDLILIKLSTPAVINARVSTL 126

Qy 173 NVSSHCPASGATKLVSGWGTGKSPQVHPFKVQLCLNISVLQKRCEDAYPRQIDDTMFC 232
Db 127 LLPSACASAGTECLISGNGTKSGSGVPYSLLOCLKAPVLSDSCKSSYPGQITGNMIV 186

Qy 233 GD-KAGRDSQGDGSGPVVCGSLQGLVSGWDYPCARPNNRPGVYTNLCKFTKWIQETIO 291
Db 187 GFLEGGKDSQGDGSGPVVCGSLQGLVSGW-YGCAQKNKPGVYTKVCNVVNIQQTIAA 245

Qy 292 NS 293
Db 246 NS 247

RESULT 12
S31779
trypsin (EC 3.4.21.4) III precursor - Atlantic salmon (fragment)
C;Species: Salmo salar (Atlantic salmon)
C;Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S66657; S31779
R;Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A;Title: Molecular cloning and characterization of anionic and cationic variants of tryp
A;Reference number: S66657; MUID:96035908; PMID:7556223
A;Accession: S66657
A;Molecule type: mRNA
A;Residues: 1-238 <MAL>
A;Cross-references: UNIPROT:P35033; EMBL:X70074; NID:g64387; PIDN:CAA49679.1; PID:g64388
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;1-7/Domain: signal sequence (fragment) #status predicted <SIG>
F;8-15/Domain: activation peptide #status predicted <APT>
F;16-238/Product: trypsin III #status predicted <MAT>
F;16-231/Domain: trypsin homology <TRY>
F;22-152,40-56,124-225,131-198,163-177,188-212/disulfide bonds: #status predicted
F;55,99,192/Active site: His, Asp, Ser #status predicted

Query Match 34.7%; Score 558.5; DB 2; Length 238;
Best Local Similarity 44.6%; Pred. No. 4.9e-42;
Matches 104; Conservative 42; Mismatches 82; Indels 5; Gaps 4;

Qy 61 DSSSRRIINGSCDMDHPTQWQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Db 10 DDEDDKIVGGYECRKNSASYQAS--LQSGYHFCGSLISSTWVWSAAHCKYSRIQVRLGE 67

Qy 121 YSLSPVYSGQOMFQGVKSIPIHPGVSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPCA 180
Db 68 HNTA-VNEGTEQFIDSVKVIMHPSYNSRNLNDMLIKLSPASLSNVSTVALPSSCAS 126

Qy 181 AGTKLVSGWGTGKSPQVHPFKVQLCLNISVLQKRCEDAYPRQIDDTMFCAGD-KAGRD 239
Db 127 SGTRCLVSGWNLISGSSSNYPDTLRCLDLPLILSSSSCNSSAYPGQITSNMFCAGFWEGKD 186

Qy 240 SCQGDGSGPVVCGSLQGLVSGWDYPCARPNNRPGVYTNLCKFTKWIQETIOAN 292
Db 187 SCQGDGSGPVVCGSLQGLVSGW-YGCAQKNKPGVYTKVCNVYNSWISSTSSNS 238

RESULT 13
TRDGC
trypsin (EC 3.4.21.4) precursor, cationic - dog
N;Alternate names: cationic trypsinogen
C;Species: Canis lupus familiaris (dog)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C;Accession: B26273

```
QY      65 SRIINGSDCDMHTQPWQAALLLRPNQLYCAGVLVHPQWLTTAAHCRKKKVRFRVLGHYSLS 124  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db      8 SRIVGGYCKEKSQPVQVAII--NEYLCCGGLIDPSSWITTAHCYSNYHYHVLLGRNNLF 64  
  
QY     125 PYVESQQMFQGVKSIPIPHGY-----SH-----PGHSNDMLIKLNRRIRPTKDVRPIN 173  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db     65 EDEPPAQRYFVN-OSFPHPDYPKFPLMRNHRTRGDGDSNDLMLHLUSEPADITDGKVVID 123  
  
QY    174 VSSHCPASCTKLVSQGWGTTSPOVHFPPVKLOCLNISLVLSQRKCEDAYPRQIDDWTMFCAG 233  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db   124 LTPEEPKVGSTCLASGSWGSKPLNWELPDLDQCWNHIILSNKEKCIAYEQKVTDLMLCAG 193  
  
QY    234 DKAGR-DSCQGSGGPVVCGNSLGQIWSWDGPCARPNGPYVTNLCKFTKIQTIOAN 292  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db   184 EMDDRKTCKDGGSLICDGLVLOGITSXGNVPCEAPNPYGYYTKLIKFTSWKEYMKEN 243  
  
RESULT 15  
JQ1472  
C:Species Rattus norvegicus (Norway rat)  
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004  
C:Accession JQ1472  
R;Kang, J.; Wiegand, U.; Mueller-Hill, B.  
Gene 110, 181-187, 1992  
A>Title: Identification of cDNAs encoding two novel rat pancreatic serine protease genes from embryonic kidney fibroblasts  
A:Reference number: JQ1471; MUID:92165057; PMID:1537555  
A:Accession: JQ1472  
A:Molecule type: mRNA  
A:Residues: 1-246 <x>  
A:Cross-references: UNIPROT:P32822; EMBL:X59013; NID:g57414; PIDN:CAA41752.1; PDB:1OZT  
A:Experimental source: pancreas  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen  
F;1-15/Domain: signal sequence #status predicted <SIG>  
F;16-24/Domain: activation peptide #status predicted <ACT>  
F;25-246/Product: trypsin V, b-form #status predicted <MAT>  
F;25-239/Domain: trypsin homology <TRY>  
F;31-160,49-65,133-233,140-206,171-185/Disulfide bonds: #status predicted  
F;64,108,200/Active site: His, Asp, Ser #status predicted
```

Query Match 33.8%; Score 544; DB 2; Length 246;
Best Local Similarity 46.2%; Pred.No. 9.7e-41;
Matches 108; Conservative 32; Mismatches 88; Indels 6; Gaps 5;

```
QY      60 SDSSSRINGSDCDMHTOPWAALLLRPNQLYCGAVLVHPQWLTTAAHCRKKKVFVRVLG 119  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db     18 TEDDDRIDRVGGTYCOBESHVPYQVS--LNAGSHTCGSSLITDQWLSAAACHYPOLQVRLG 75  
  
QY    120 HYSISLPYESGOQMFGVKISIPHGPSHPCHSNDMLIKLNRRIRPTKDVRPINVSSHCP 179  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db     76 EHNIYEI-EGAEGFDIAAKMIHPDYDKMTVDINDIMLKLSPATLNSKVSTIPTLPQYCP 134  
  
QY    180 SAGTKLVSWGWTTSPOVHFPPVKLOCLNISLVLSQRKCEDAYPRQIDDWTMFCAGD-KAGR 238  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db   135 TAGTECLVSGWGLKF-GFESPSPVLQCLDAFLSDSVCHKAYPRQITTNNMFLCGFLLEGKG 193  
  
QY    239 DSCQSGSGGPVVCGNSLGQIWSWDGPCARPNGPYVTNLCKFTKIQTIOAN 292  
       ||| : |::||| :: |::||| :: |::||| :: |::||| :: |::||| :: |::||  
Db   194 DSCQYDGGPVCNGEVIGVSGMD-GCALEGRPYYTVKYCNLYNNIQOTVAAN 246
```

Search completed: July 4, 2005, 15:23:03
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 15:07:33 ; Search time 71 seconds
(without alignments)
2113.228 Million cell updates/sec

Title: US-09-936-271C-14

Perfect score: 1608

Sequence: 1 MATARPPMWWLCAITALL.....VVTNLCKFTKWIQETIQANS 293

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_Q3:*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	293	1 KLK5 HUMAN	Q9Y337 homo sapien
2	1475	91.7	269	2 Q8IU55	Q8IU55 homo sapien
3	1158.5	72.0	293	2 Q9DI40	Q9DI40 mus musculus
4	1074	66.8	204	2 Q6S9W8	Q6S9W8 homo sapien
5	683	42.5	276	2 Q8CGR6	Q8CGR6 mus musculus
6	676	42.0	254	2 Q9XSN6	Q9XSN6 sus scrofa
7	671	41.7	254	1 KLK4 HUMAN	Q9Y5K2 homo sapien
8	654	40.7	246	2 Q6P3Z0	Q6P3Z0 mus musculus
9	650.5	40.5	249	2 Q9QYN4	Q9QYN4 mus musculus
10	650.5	40.5	250	1 KLBK HUMAN	Q9UBX7 homo sapien
11	650.5	40.5	250	2 Q632F2	Q632F2 rattus norv
12	650.5	40.5	276	2 Q9QYN3	Q9QYN3 m hipposas
13	643.5	40.0	260	2 Q8TW69	Q8TW69 homo sapien
14	641.5	39.9	256	2 Q6IE12	Q6IE12 rattus norv
15	641.5	39.9	260	1 KLBK HUMAN	Q60259 homo sapien
16	638	39.7	255	2 Q9JIS2	Q9JIS2 mus musculus
17	636.5	39.6	277	1 KLBK HUMAN	Q9UKR3 homo sapien
18	634	39.4	255	2 Q9Z0M1	Q9Z0M1 mus musculus
19	629	39.1	275	2 Q8IXD7	Q8IXD7 homo sapien
20	627.5	39.0	260	1 NRPN RAT	Q88780 rattus norv
21	627	39.0	260	1 NRPN MOUSE	Q61955 mus musculus
22	623.5	38.8	251	2 Q6B089	Q6B089 homo sapien
23	620.5	38.6	251	1 KLBK HUMAN	Q9P0G3 homo sapien
24	612.5	38.1	242	2 Q8OV84	Q8OV84 mus musculus
25	612.5	38.1	250	2 Q8CGR5	Q8CGR5 mus musculus
26	612.5	38.1	253	1 KLBK HUMAN	P49862 homo sapien
27	601.5	37.4	253	2 Q8NSN9	Q8NSN9 homo sapien
28	599.5	37.3	250	1 KLBK HUMAN	Q9UKQ9 homo sapien
29	594	36.9	249	2 Q9IVE3	Q9IVE3 m thymopsin
30	586.5	36.5	247	2 Q9CPN9	Q9CPN9 m mus muscu
31	586	36.4	246	1 TRY1 RAT	P00762 rattus norv

32	585.5	36.4	258	2 Q63ZP4	Q63ZP4 rattus norv
33	582.5	36.2	247	1 TRY3 RAT	P08426 rattus norv
34	580.5	36.1	247	2 Q9D7Y7	Q9D7Y7 mus musculus
35	578.5	36.0	255	2 TRYJG6	P71196 saguinus oe
36	578	35.9	247	1 TRY2 CANFA	P06872 canis famil
37	574.5	35.7	205	2 Q96JE2	Q96JE2 homo sapien
38	574	35.7	234	2 Q9CV76	Q9CV76 mus musculus
39	574	35.7	243	1 TRY1 BOVIN	P00760 bos taurus
40	574	35.7	255	2 Q6GNJ2	Q6GNJ2 xenopus lae
41	573.5	35.7	237	2 Q6GIJ5	Q6GIJ5 struthio ca
42	573	35.6	244	1 TRY2 XENLA	P70059 xenopus lae
43	573	35.6	248	2 Q7SZT1	Q7SZT1 xenopus lae
44	569.5	35.4	248	1 TRY3 CHICK	Q90629 gallus gall
45	569.5	35.4	255	2 Q6ISIO	Q6ISIO homo sapien

ALIGNMENTS

RESULT 1

ID	KLK5_HUMAN	HUMAN	STANDARD;	PRT;	293 AA.
AC	Q9Y337; Q9HBG8;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	25-OCT-2004 (Rel. 45, Last annotation update)				
DE	Kallikrein 5 precursor (EC 3.4.21.-) (Stratum corneum tryptic enzyme)				
DE	(Kallikrein-like protein 2) (KLK-L2) (UNQ570/PRO1132).				
GN	Name=KLK5; Synonyms=SCTE;				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Stratum Corneum;				
RA	MEDLINE=99445563; PubMed=10514489; DOI=10.1074/jbc.274.42.30033;				
RA	Brattand M., Egelrud T.;				
RT	"Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation."				
RL	J. Biol. Chem. 274:30033-30040(1999).				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20118156; PubMed=10652563;				
RA	Yousef G.M., Luo L.-Y., Diamandis E.P.;				
RT	"Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4."				
RL	Anticancer Res. 19:2843-2852(1999).				
RP	SEQUENCE FROM N.A.				
RX	PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;				
RA	Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,				
RA	Moss P., Paepker B., Wang K.;				
RT	"Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region."				
RL	Gene 257:119-130(2000).				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;				
RA	Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,				
RA	Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,				
RA	Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,				
RA	Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,				
RA	Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,				
RA	Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,				
RA	Vandlen R., Watanabe C., Wiedand D., Woods K., Xie W.-H., Yansura P.,				
RA	Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,				
RA	Godowski P., Gray A.;				
RT	"The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment."				
RL	Genome Res. 13:2265-2270(2003).				

[5]
SEQUENCE FROM N.A.
TISSUE=Ovary;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschuler R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: May be involved in desquamation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in skin, breast, brain and testis.
CC -!- SIMILARITY: Belongs to the peptidase S1 family, Kallikrein
subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF168768; AAF03101.1; -;
DR EMBL; AF135028; AAD26429.1; -;
DR EMBL; AF243527; AAG33358.1; -;
DR EMBL; AY359010; AAQ89369.1; -;
DR EMBL; BC008036; AAH08036.1; -;
DR HSSP; P00760; 1EZX.
DR MEROPS; S01.017; -;
DR Genew; HGNC:6366; KLK5.
DR MW; 605643; -;
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0008544; P:epidermal differentiation; TAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolyase; Serine protease; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 293 Kallikrein 5.
FT ACT_SITE 108 108 Charge relay system (By similarity).
FT ACT_SITE 153 153 Charge relay system (By similarity).
FT ACT_SITE 245 245 Charge relay system (By similarity).
FT DISULFID 73 206 By similarity.
FT DISULFID 93 109 By similarity.
FT DISULFID 178 279 By similarity.
FT DISULFID 185 251 By similarity.
FT DISULFID 217 231 By similarity.
FT DISULFID 241 266 By similarity.
FT CARBOHYD 69 69 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 173 173 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 208 208 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 252 252 N-linked (GlcNAc...) (Potential).
FT CONFLICT 25 56 Missing (in Ref. 3).
SQ SEQUENCE 293 AA; 32020 MW; D92C92F5609E5946 CRC64;
Query Match 100.0%; Score 1608; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 5.6e-139;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATATPPMMVLCALITALLIGVTEHVLANNVSCDHPSTVPSGNSQDLGAGAGEARS 60
DB 1 MATATPPMMVLCALITALLIGVTEHVLANNVSCDHPSTVPSGNSQDLGAGAGEARS 60
QY 61 DSSSRRIINGSDCDMHTQPMQALLLPNOLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
DB 61 DSSSRRIINGSDCDMHTQPMQALLLPNOLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120
QY 121 YSLSPVYESGOQMFQGVKSIPIHGYSHPHGSHNDLMLIKLNRRIRPTKDVPRINVSCHPS 180
DB 121 YSLSPVYESGOQMFQGVKSIPIHGYSHPHGSHNDLMLIKLNRRIRPTKDVPRINVSCHPS 180
QY 181 AGTKCLVSGHGTGTSPOVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGHGTGTSPOVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPNNRPGVYTNLCCKFTKWIQETIQANS 293
DB 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPNNRPGVYTNLCCKFTKWIQETIQANS 293
RESULT 2
Q8IU55 PRELIMINARY; PRT; 269 AA.
ID Q8IU55
AC Q8IU55; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Kallikrein 5 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovarian cancer;
RX MEDLINE=22623266; PubMed=12738725;
RA Dong Y., Kaushal A., Brattsand M., Nicklin J., Clements J.A.;
RT "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer
produces novel variants with potential as cancer biomarkers.";
RL Clin. Cancer Res. 9:1710-1720(2003).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF435980; AAN63606.1; -;
DR EMBL; AF435981; AAN63607.1; -;
DR HSSP; P00760; 1EZX.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 29226 MW; D3C1E1FBEEA3634C CRC64;
Query Match 91.7%; Score 1475; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 7.7e-127;

```
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATATPPMMVLCALITALLGLVTEHVLANDVSCDHPSNTVPVSGNODLGAGAGEDARS 60
DB 1 MATATPPMMVLCALITALLGLVTEHVLANDVSCDHPSNTVPVSGNODLGAGAGEDARS 60
QY 61 DSSSRIRNGSDCDMHTQPMQAAALLRPNOYLCGAVLVHPQWLLTAACHCKKVFVRVLRGH 120
DB 61 DSSSRIRNGSDCDMHTQPMQAAALLRPNOYLCGAVLVHPQWLLTAACHCKKVFVRVLRGH 120
QY 121 YSLSPVYESGQGMFOGVKSIPIHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHCPS 180
DB 121 YSLSPVYESGQGMFOGVKSIPIHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHCPS 180
QY 191 AGTKCLVSGWGTGTSKSPQVHFPRKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 191 AGTKCLVSGWGTGTSKSPQVHFPRKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQGSDDGGPVVCGNSLQGLVSGWDYPCARP 269
DB 241 CQGSDDGGPVVCGNSLQGLVSGWDYPCARP 269

RESULT 3
Q9D140 PRELIMINARY; PRT; 293 AA.
ID Q9D140
AC Q9D140;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:110030019 product:weakly similar to KALLIKREIN 5 (EC
DE 3.4.21.-) (STRATUM CORNEUM TRYPTIC ENZYME) (KALLIKREIN-LIKE PROTEIN 2)
DE (KLK-L2).
GN Name=1110030019Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body; DOI=10.1016/S0076-6879(99)03004-9;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
```

```
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumi T., Tashiro H., Itoh M.,
RA Sumi N., Ishi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AK003996; BAB23113.1; -
DR HSSP; P00760; 1BZX.
DR MEROPS; S01.418; -
DR MGB; MGI:1915916; 1110030019Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR003003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 293 AA; 31908 MW; ED1F45D8226FE911 CRC64;

Query Match 72.0%; Score 1158.5; DB 2; Length 293;
Best Local Similarity 69.7%; Pred. No. 8.6e-98;
Matches 205; Conservative 41; Mismatches 45; Indels 3; Gaps 2;
QY 1 MATATPPMMVLCALITALLGLVTEHVLANDVSCDHPSNTVPVSGNODLG--AGAGEDA 58
DB 1 MARTGHPWKWAMATLITTLVLGVSEPVLAGDVSSCDNPSGTEPSGTNRDLSTDSKSGDT 60
QY 59 RSDSSSRIRNGSDCDMHTQPMQAAALLRPNOYLCGAVLVHPQWLLTAACHCKKVFVRVRL 118
DB 61 RS-DSSSRIRVNGSDCQKDAQPMQWQALLGLPNKLYCGAVLISPMWLLTAACHCKKVFVRVRL 119
QY 119 GHYSLSPVYESGQGMFOGVKSIPIHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHC 178
DB 120 GHHSMSPVYESGQGMFOGVKSIPIHPGYSHPGHNDMLIKLNRRIRPTKDVPRINVSCHC 179
QY 179 PSAGTKCLVSGWGTGTSKSPQVHFPRKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGR 238
DB 180 ATEGTRCMVSCWGTGTTSSSHNNFPKVLQCLNITLVSEERCKNSYFQIDKTMFCAGDEGR 239
QY 239 DSCGSDGGPVVCGNSLQGLVSGWDYPCARPNRPGVYTNLCKFKTKWIOETQAN 292
DB 240 DSCGSDGGPVVCGNSLQGLVSGWDYPCARPNRPGVYTNLCEFKVKWIKDTMNSN 293

RESULT 4
Q6S9W8 PRELIMINARY; PRT; 204 AA.
ID Q6S9W8
AC Q6S9W8;
```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Kallikrein 5 isoform 3 preproprotein.
GN Name=KUK5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBT_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Michael I.P., Kurlender L., Du D.C., Diamandis E.P.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY461805; AAR23814.1; -.
DR HSP; P36368; IAO5.
DR GO; GO:0004283; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 204 AA; 22100 MW; C9FFDA022246BD52 CRC64;

Query Match 66.8%; Score 1074; DB 2; Length 204;
Best Local Similarity 98.5%; Pred. No. 3.1e-90;
Matches 198; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATATPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPSGNSQDLGAGEDARS 60
Db 1 MATATPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPSGNSQDLGAGEDARS 60

Qy 61 DSSSRIRNGSDCMHTOPWQAALLRNQLYCGAVLHPQWLLTAHCRKKVFRVLGH 120
Db 61 DSSSRIRNGSDCMHTOPWQAALLRNQLYCGAVLHPQWLLTAHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMGFGVKSIPHPGSHPGSHNDLMLIKLNRIRPTKDVRLPINVSSHCP 180
Db 121 YSLSPVYESGQMGFGVKSIPHPGSHPGSHNDLMLIKLNRIRPTKDVRLPINVSSHCP 180

Qy 181 AGTKCLVSGWGTGTSKSPQVHFP 201
Db 181 AGTKCLVSGWGTGTSKSPQGECP 201

RESULT 5
Q8CGR6 ID Q8CGR6 PRELIMINARY; PRT; 276 AA.
AC Q8CGR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Glandular kallikrein KUK13.
GN Name=Kuk13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBT_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=2325484; PubMed=12437987; DOI=10.1016/S0006-291X(02)02629-3;
RA Olsson A.Y., Lundwall A.;
RT "Organization and evolution of the glandular kallikrein locus in Mus musculus.";
RL Biochem. Biophys. Res. Commun. 299:305-311(2002).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RA Adams M., Mural R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY152432; AAN7820.1; -.
DR HSP; P00760; IEZX.
DR MEROPS; S01.306; -.
DR MG; MGI:95292; K1k13.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 276 AA; 30360 MW; 33E189C67492FDCA CRC64;

Query Match 42.5%; Score 683; DB 2; Length 276;
Best Local Similarity 46.2%; Pred. No. 3e-54;
Matches 133; Conservative 50; Mismatches 75; Indels 30; Gaps 8;

Qy 9 MWVLCALITALLGVTEHVLANNVSCDHPSTNTVPSGNSQDLGAGEDARSDDSSRII 68
Db 1 MWPLVATIACLTALSE-----GISRDYK-ILNGTN-----GTSGLP 38

Qy 69 NGSDCMHTOPWQAALLRNQLYCGAVLHPQWLLTAHCRKKVFRVLGHYSLSPYE 128
Db 39 GGYTCLPHSQPWAALLIR-GRLLCGVLVHPKWLTAHCRKDGVTVHLGKHALGRV-E 96

Qy 129 SGQMGFGVKSIPHPG-----SHPGSHNDLMLIKLNRIRPTKDVRLPINVSS-HCPSAGT 183
Db 97 NGEQMEVRSIHPPEYQVPTHLNHDHIMLELSPVQLSSHVRLTKLSADCLPTGT 156

Qy 184 KCLVSGWGTGTSKSPQVHFPKVLQCLINISVLQKRCEDAYPRQIDTFMFCAGDK-AGRDSQC 242
Db 157 CCRVSGWGTGTSKSPQVNYPKTLQCANIELRSDDECRQVYPGKITANMLCAGTKEGKDSCE 216

Qy 243 GDSGGPVVCSGSLQGLVSGWDYPCARNPGRPGVYTNLCCKFTKWIQETIQ 290
Db 217 GDSGGLICNGKLYIGISWGDYPCGQPNRPGVYTRVSKYLRIEIR 264

RESULT 6
Q9XSN6 ID Q9XSN6 PRELIMINARY; PRT; 254 AA.
AC Q9XSN6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Enamel matrix serine proteinase 1 precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBT_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98126310; PubMed=9465170;
RA Simmer J.P., Fukae M., Tanabe T., Yamakoshi Y., Uchida T., Xue J.,
RA Margolis H.C., Shimizu M., Dehart B.C., Hu C.-C., Bartlett J.D.;
RT "Purification, characterization, and cloning of enamel matrix serine proteinase 1.";
RL J. Dent. Res. 77:377-386(1998).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; U76256; AAB94638.1; -.
DR HSP; P00760; IEZX.
DR MEROPS; S01.251; -.
DR GO; GO:0004283; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

```

DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0020; TRYP SEC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 24 Potential.
FT CHAIN 31 254 enamel matrix serine proteinase 1.
SQ SEQUENCE 254 AA; 27235 MW; FD40EP85664406F1 CRC64;

Query Match 42.0%; Score 676; DB 2; Length 254;
Best Local Similarity 47.4%; Pred. No. 1.2e-53;
Matches 139; Conservative 35; Mismatches 79; Indels 40; Gaps 6;

QY 1 MATARPPMMWVLCALITALLGVTEHVLNNDVSCDHPSTNTPVPSGNSQDLGAGAGEDARS 60
DB 1 MTRAGNPGWEL-----GYLLIGIT-----GSLAFINGG----- 29

QY 61 DSSSRIRSGDCDMHTQWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120
DB 30 -----HIINGEDCNPHSQPQAALFLE-DDFCGGVLVHPQWLSAAHCFQNSVYIGLGL 83

QY 121 YSLSPVSVESGQMPGQGVKSIHPGYSYHPCGHNLDMLIKLNRIRPTKDVRIINVSCHPS 180
DB 84 HNLPEQEPGQSMQMEASLSIHPNEYNEPMSANDMLIKLRESVLSLTVRINISVVSQCPT 143

QY 181 AGTKLVSGMGWTKSPQVHFVKVQLQNLISVLRSOKRCEDAYPRQIDDTMPCA-GDKAGRD 239
DB 144 PGDSCLVSGWGLASGLI--PQVLCVNISVASEVCKARGVYHPSPNFCAGGQDQKD 201

QY 240 SCQSDSGPVVNCVSGLSQGLVSGWYPCARPNGVYTNLCKFTKWIQTIQAN 292
DB 202 SCHGDSGGPLICNGSLQGLVSGWYPCARPNGVYTNLCKFTDWTQTIQAS 254

RESULT 7
KLK4 HUMAN
ID_KLK4 HUMAN STANDARD; PRT; 254 AA.
AC Q9Y5K2; O9GZL6; Q9UBJ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 25-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Kallikrein 4 precursor (EC 3.4.21.-) (Protease) (Kallikrein-like protein 1) (KLK-L1) (Enamel matrix serine proteinase 1).
GN Name=KLK4; Synonyms=EMSP1, PRSS17, PSTS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99179024; PubMed=10077646; DOI=10.1073/pnas.96.6.3114;
RA Nelson P.S., Gan L., Ferguson C., Moss P., Gellinas R., Hood L., Wang K.;
RT "Molecular cloning and characterization of prostate, an androgen-regulated serine protease with prostate-restricted expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3114-3119 (1999).
RP SEQUENCE FROM N.A.
RX MEDLINE=99413477; PubMed=10485467;
RA Yousef G.M., Obiezu C.V., Luo L.-Y., Black M.H., Diamandis E.P.;
RT "Protease/KLK-L1 is a new member of the human kallikrein gene family, is expressed in prostate and breast tissues, and is hormonally regulated.";
RL Cancer Res. 59:4252-4256 (1999).
RP [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99367447; PubMed=10438493; DOI=10.1074/jbc.274.33.23210;
RA Stephenson S.A., Verity K., Ashworth L.K., Clements J.A.;

RT "Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3-13.4.";
RL J. Biol. Chem. 274:23210-23214 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepers B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";
RL Gene 257:119-130 (2000).
RN [5]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION
RX MEDLINE=20323211; PubMed=10863090; DOI=10.1016/S0378-1119(00)00203-1;
RA Hu J.-C., Zhang C., Sun X., Cao X., Ryu O., Simmer J.P.;
RT "Characterization of the mouse and human PRSS17 genes, their relationship to other serine proteases, and the expression of PRSS17 in developing mouse incisors.";
RL Gene 251:1-8 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RX TISSUE=Prostatic carcinoma;
RA Korkmaz K.S., Korkmaz C.G., Saatcioglu F.;
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RX PubMed=15057824; DOI=10.1038/nature02399;
RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J., Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M., Aerts A., Alther M., Ashworth L., Bajorek E., Black S., Branscomb E., Caepeel S., Carrano A.V., Caoille C., Chan Y.M., Christensen M., Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Deter J.C., Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M., Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T., Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H., Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S., Malfatti S., Martinez D., McCreedy P.M., Medina C., Morgan J., Nelson K., Nolan M., Ovcharenko I., Pittluck S., Poillard M., Popkie A.P., Predki P., Quan Q., Ramirez L., Rash S., Retterer J., Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D., Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A., Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I., Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E., Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M., Rubin E.M., Lucas S.M.;
RT "The DNA sequence and biology of human chromosome 19.";
RL Nature 428:529-535 (2004).
RN [8]
RP SEQUENCE OF 22-254 FROM N.A.
RX Simmer J.P., Ryu O.H., Qian Q., Zhang C., Cao X., Sun X., Hu C.-C.;
RT "Cloning and characterization of a cDNA encoding human EMSP1.";
RL (In) Goldberg M. (eds.);
RP Chemistry and biology of mineralized tissues, pp.1-1, American Academy of Orthopaedic Surgeons, Vitell (2000).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in prostate.
CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF113140; AAD21580.1; -
DR EMBL; AF113141; AAD21581.1; -
DR EMBL; AF135023; AAD26424.2; -
DR EMBL; AF148532; AAD38019.1; -
DR EMBL; AF243527; AAG33357.1; -


```

[7]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Testis;
MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603999;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore S.I., Wang J., Heide F.,
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Maillu D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: Possible multifunctional protease. Efficiently cleaves
bz-Phe-Arg-4-methylcoumaryl-7-amide, a kallikrein substrate, and
weakly cleaves other substrates for kallikrein and trypsin.
-!- SUBCELLULAR LOCATION: Secreted.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q9UBX7-1; Sequence=Displayed;
Name=2;
IsoId=Q9UBX7-2; Sequence=VSP 005402;
-!- TISSUE SPECIFICITY: Expressed in brain, skin and prostate. Isoform
1 is expressed preferentially in brain; isoform 2 in prostate.
-!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AB012917; BAA33404.1; AUT_INIT.
EMBL; AB013730; BAA88713.1; -.
EMBL; AB041036; BAA96797.1; -.
EMBL; AF164623; AAD47815.1; -.
EMBL; AF243527; AAG33364.1; -.
EMBL; AY359014; AAG99373.1; -.
EMBL; AC011473; AAG23257.1; -.
EMBL; BC022068; AAH22068.1; -.
HSP; P00760; 1EXZ.
MEROPS; S01.257; -.
Genew; HGNC:6359; KUK11.
H-invDB; HIX0015375; -.
MIM; 604434; -.
GO; GO:0008236; F:serine-type peptidase activity; TAS.
InterPro; IPR009003; Pept_Ser_Cys.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00089; Trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Alternative splicing; Glycoprotein; Hydrolase; Serine protease;
Signal; Zymogen.
SIGNAL 1 18 Potential.
PROPEP 19 21 Activation peptide (Potential).
```

```

FT CHAIN 22 250 Kallikrein 11.
FT ACT SITE 62 62 Charge relay system (By similarity).
FT ACT SITE 110 110 Charge relay system (By similarity).
FT ACT SITE 203 203 Charge relay system (By similarity).
FT DISULFID 28 163 By similarity.
FT DISULFID 47 63 By similarity.
FT DISULFID 135 237 By similarity.
FT DISULFID 142 209 By similarity.
FT DISULFID 174 188 By similarity.
FT DISULFID 199 224 By similarity.
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 165 165 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 181 181 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 210 210 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 1 1 M -> MORLRLRWKSSGRGLTAAKPGARSSPLQAM
(in isoform 2).
FT FT /FTId=VSP_005402
SQ SEQUENCE 250 AA; 27466 MW; 192D910BECDC7A56 CRC64;

Query Match 40.5%; Score 650.5; DB 1; Length 250;
Best Local Similarity 50.6%; Pred. No. 2.5e-51;
Matches 118; Conservative 41; Mismatches 67; Indels 7; Gaps 4;

Qy 65 SRIINGSDCDMHTQFWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGHYSLS 124
Db 20 TRIIKGFCKPHSQPQWQA-LFEKTRLLCGATLAPRWLLTAARCLPRYIVHLGQHNLQ 78
Qy 125 PVYESGOQMFQGVKSIHPGVS---HPGHSNDLMLKLNRRIRPTKDVRLINVSCHPCS 180
Db 79 K-EGCEQTATATESFPHPGFNSLPNKHNRNDLMLVKMASPVSIITWAVRLTUSRCVT 137
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSQKRCEDAYPRQIDDTMFAG-DKAGRD 239
Db 138 AGTSLISGWSSTSPQLRLPHTLRCAITIIIEHKCNAYPGNITDTWVCASVQEGSKD 197
Qy 240 SCQSDSGPVVNGSLGSLGSLGSDGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYP 292
Db 198 SCQSDSGPLVCSGLOGIISWGQDPCAITRKPGVYTKVCKYVDWIOBTMKN 250

RESULT 11
Q63ZF2 Q63ZF2 PRELIMINARY; PRT; 250 AA.
AC Q63ZF2; 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Glandular kallikrein 11 precursor.
GN Name=Klik1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15203212; DOI=10.1016/j.ygeno.2004.01.009;
RA Olsson A.Y., Lilja H., Lundwall A.;
RT "Taxon-specific evolution of glandular kallikrein genes and
RT identification of a progenitor of prostate-specific antigen.";
RL Genomics 84:147-156(2004).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK001364; DAA05600.1; -.
KW Signal.
FT SIGNAL 1 18 Potential.
FT CHAIN 22 250 glandular kallikrein 11.
SQ SEQUENCE 250 AA; 27630 MW; 247B29D3F0DEF8F1 CRC64;

Query Match 40.5%; Score 650.5; DB 2; Length 250;
Best Local Similarity 50.6%; Pred. No. 2.5e-51;
Matches 118; Conservative 40; Mismatches 68; Indels 7; Gaps 4;

Qy 65 SRIINGSDCDMHTQFWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGHYSLS 124
```

```
Db 20 TRIKGYECRPHSQPQVA-LFQKTRLLCGATLIAPKWLITAAHCKRKHVYLLGEHNL 78
Qy 125 PYESGQOMFGQVKSIPHPGYS----HPGSHNDMLIKLNRRIRPTKDVPRINSSHCP 180
Db 79 KT-DGCEQRMMATESFPFHPGNSLPNKHNDIMLVKMSSPAFITRAVRPLTSSLCVT 137
Qy 181 AGTKCLVSGWGTTSQPQVHFPKVLQCLNLSVLSQKRCEDAYPRQIDDDTMFCAG-DK 239
Db 138 AGTSLISGWGTTSSPQLRPHSLRCANVSIIHGKECERAYPGNITDMLCASVRKEGD 197
Qy 240 SCQDGSQGVVNCGLSVGMDYPCARPNGPYTYTLNCKRPTKWIQTIQAN 292
Db 198 SCQDGSQGVVNCGLSVGMDYPCARPNGPYTYTLNCKRPTKWIQTIQAN 250

RESULT 12
Q9QYN3 PRELIMINARY; PRT; 276 AA.
AC Q9QYN3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hipostatinsin prostate type (Mus musculus adult male tongue cDNA, RIKEN
DE full-length enriched library, clone:2310015108 product:protease,
DE serine, 20, full insert sequence) (Mus musculus adult male tongue
DE cDNA, RIKEN full-length enriched library, clone:2310040F07
DE product:protease, serine, 20, full insert sequence).
GN Name=2310015108Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=20525460; PubMed=11072088; DOI=10.1016/S0167-4781(00)00206-2;
RA Mitsui S., Okui A., Koninami K., Uemura H., Yamaguchi N.;
RT "cDNA cloning and tissue-specific splicing variants of mouse
RT hippocasin/TLSP (PRSS20).";
RL Biochim. Biophys. Acta 1494:206-210(2000).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=92729253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
```

```
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20530913; PubMed=11078661; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AB016227; BAA36955.1; -
DR EMBL; AK009360; BAB26241.2; -
DR EMBL; AK009720; BAB26461.2; -
DR HSP; P00760; IEZX.
DR MEROPS; S01.257; -.
DR MG; MG1:1923977; 2310015108Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 276 AA; 30753 MW; 90BDC03A8AB178D6 CRC64;

Query Match 40.5%; Score 650.5; DB 2; Length 276;
Best Local Similarity 49.8%; Pred. No. 2.8e-51;
Matches 116; Conservative 42; Mismatches 68; Indels 7; Gaps 4;

Qy 65 SRIINGSDCDHMTQPMQAAILLRPNQLYCGAVLVHPQWLLTAAHCKRKHVYLLGEHNL 124
Db 46 TRIKGYECRPHSQPQVA-LFQKTRLLCGATLIAPKWLITAAHCKRKHVYLLGEHNL 104

Qy 125 PYESGQOMFGQVKSIPHPGYS----HPGSHNDMLIKLNRRIRPTKDVPRINSSHCP 180
Db 105 KT-DGCEQRMMATESFPFHPGNSLPNKHNDIMLVKMSSPAFITRAVRPLTSSLCVT 163

Qy 181 AGTKCLVSGWGTTSQPQVHFPKVLQCLNLSVLSQKRCEDAYPRQIDDDTMFCAG-DK 239
Db 164 AGTSLISGWGTTSSPQLRPHSLRCANVSIIHGKECERAYPGNITDMLCASVRKEGD 223

Qy 240 SCQDGSQGVVNCGLSVGMDYPCARPNGPYTYTLNCKRPTKWIQTIQAN 292
Db 224 SCQDGSQGVVNCGLSVGMDYPCARPNGPYTYTLNCKRPTKWIQTIQAN 276
```

RESULT 13

ID	Q8IW69		PRELIMINARY;	PRT; 260 AA.
AC	Q8IW69			
DT	01-MAR-2003	(T=EMBLrel. 23, Created)		
DT	01-MAR-2003	(T=EMBLrel. 23, Last sequence update)		
DT	01-MAR-2004	(T=EMBLrel. 26, Last annotation update)		
DE	Kallikrein 8, isoform 1 preproprotein.			
GN	Name=KLK8;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Etheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCB_I_TaxID=9606;			
RN	[1]			
SEQUENCE FROM N.A.				
TISSUE=Brain;				
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA Klausner R.D., Collins F.S., Wagner K.J., Shenmen C.M., Schuler G.D.,				
RA Altschul S.F., Ziegler B., Buettow K.H., Schaefer C.F., Bhut N.K.,				
RA Hopkins R.E., Jordan H., Moore T., Max S.I., Wang X.J., Hsieh F.,				
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,				
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA Brownstein M.J., Usdin T.B., Toshioyuki S., Carninci P., Prange C.,				
RA Rana S.S., Loquellano N.A., Peters G.U., Abramson R.D., Mullany S.J.,				
RA Bosak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA Richards S., Worley K.E., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,				
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
FAhey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,				
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA Krzywinski M.I., Skalkale U., Smalhus D.E., Schmerch A., Schein J.E.,				
RA Jones S.J., Marra M.A.;				
RT "Generation and initial analysis of more than 15,000 full-length human				
and mouse cDNA sequences.";				
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). [2]				
SEQUENCE FROM N.A.				
TISSUE=Brain;				
RA Strausberg R.;				
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.				
CC -I- SIMILARITY: Belongs to peptidase family S1.				
DR EMBL; BC040887; AH40887.l; --				
DR HSSP; P00760; IEXZ.				
DR GO; GO:0004263; F:Chymotrypsin activity; IEA.				
DR GO; GO:0008233; F:Peptide activity; IEA.				
DR GO; GO:0004295; F:Tryptsin activity; IEA.				
DR GO; GO:0006508; P:Proteolysis and peptidolysis; IEA.				
DR InterPro; IPRO01254; Peptidase_S1.				
DR InterPro; IPRO01314; Peptidase_SIA.				
DR InterPro; IPRO09003; Pept Ser_Cys.				
PFam; PF00089; Trypsin; 1.				
PRINTS; PR00722; CHYMOTRYPSIN.				
SMART; SM00020; TRYD_SPCL; 1.				
PROSITE; PS00240; TRYP SIN DOM; 1.				
PROSITE; PS00134; TRYP SIN HIS; UNKNOWN_1.				
PROSITE; PS00135; TRYP SIN SER; 1.				
KW Hydrolase; Protease; Serine protease.				
SQ SEQUENCE 260 AA; 28090 MW; EF5934EB96295660 CRC64;				
Query Match	40.0%; Score 643.5; DB 2; Length 260;			
Best Local Similarity	42.6%; Pred. No. 1.2e-50;			
Matches 126; Conservative 43; Mismatches 80; Indels 47; Gaps 7;				
QY 1 MATARP-----PMWWVLCALITALLGVTEHVLANNDVSCDHSPNTVPSSGSLDGAGAGE 56				
DB 1 MGRRPRPAAKTWFL-----LLLG-----GAWAH 25				
57 DARSDSSRIINGSDCDMHTQPMAQLLRPNOLYCGLAVHPQWLMTAAHCCKVKFRV 116				
26 GRADED---KVIGHGECOPHSFOWAA-LFOGOILLCGVLVGNNLVLTAAHCKPKVT 81				

OY 181 AGTKCLVSGWGTGKSPVHFVKVLQCLNLSVLSOKRCEDAYPROIDDTMFAGDKAGR-D 239
DB 146 PGDTCLVSGWRLNGKLI--PSLQCVNLVASEETCELLLDVPYHLSNFCAGGGPRKD 203
OY 240 SCQSDSGPVCVNGSLQGLVSGWDPYPCARPNRPVYTNLCKFTKWIOBTIQ 290
DB 204 TCNGDSGGPIVCNRSLOGLVSGWGECQCGIPSPVYTNLCKFTNWIWTIQ 254

RESULT 15
KLX8 HUMAN STANDARD; PRT; 260 AA.
AC O60259; Q9HCB3; Q9U1L9; Q9UQ47;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Neuropilin precursor (EC 3.4.21.-) (NP) (kallikrein 8) (Ovasin) (Serine
DE protease TADG-14) (tumor-associated differentially expressed gene-14
DE protein) (UNQ283/PRO322).
GN Name=KLK8; Synonyms=NRPN, PRSS19, TADG14;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hippocampus;
RX MEDLINE=98372070; PubMed=9714609; DOI=10.1016/S0378-1119(98)00232-7;
RA Yoshida S., Taniguchi M., Hirata A., Shiosaka S.;
RT "Sequence analysis and expression of human neuropilin cDNA and gene.";
RL Gene 213:9-16(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=99203457; PubMed=10102990;
RA Mitsui S., Tsuruoka N., Yamashiro K., Nakazato H., Yamaguchi N.;
RT "A novel form of human neuropilin, a brain-related serine protease, is
RT generated by alternative splicing and is expressed preferentially in
RT human adult brain.";
RL Eur. J. Biochem. 260:627-634(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Ovary;
RX MEDLINE=99413504; PubMed=10485494;
RA Underwood L.J., Tanimoto H., Wang Y., Shigemasa K., Parmley T.H.,
RA O'Brien T.J.;
RT "Cloning of tumor-associated differentially expressed gene-14, a novel
RT serine protease overexpressed by ovarian carcinoma.";
RL Cancer Res. 59:4435-4439(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Gan L., Gelin R., Gown A.M., Moss P., Smith R., Wang K.;
RT "Molecular cloning and characterization of a novel serine protease,
RT ovasin, a potential molecular marker for ovarian carcinomas.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA Moss P., Paepel B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.-H., Yansura D.,

Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
Godowski P., Gray A.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RN Genome Res. 13:2265-2270(2003).
RN [7]
RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 1).
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Gaines J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: Suggested to be involved in kindling epileptogenesis and
CC hippocampal plasticity.
CC 1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O60259-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O60259-2; Sequence=VSP_005401;
CC 1- TISSUE SPECIFICITY: Isoform 1 is predominantly expressed in the
CC pancreas while isoform 2 is expressed in adult brain and
CC hippocampus. Both forms are also found in fetal brain and
CC placenta. Not detected in kidney, spleen, liver and lung.
CC 1- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB009849; BAA28673.1; -
CC EMBL; AB012761; BAA28676.1; -
CC EMBL; AB010780; BAA88684.1; -
CC EMBL; AB008390; BAA82665.1; -
CC EMBL; AB008927; BAA82666.1; -
CC EMBL; AF055982; AAD56050.1; -
CC EMBL; AF095742; AAD25979.1; -
CC EMBL; AF095743; AAD29574.1; -
CC EMBL; AF243527; AAG33361.1; -
CC EMBL; AY359036; AAQ89395.1; -
CC EMBL; AC011473; AAG23254.1; -
CC HSSP; P00760; 1EZX.
CC MEROPS; S01.244; -
CC Gene; HGNC:6369; KLK8.
CC MW; 605644; -
CC GO; GO:0008236; P:serine-type peptidase activity; TAS.
CC GO; GO:0007339; P:neurogenesis; TAS.
CC InterPro; IPR009003; Pept_Ser_Cys.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; Trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Alternative splicing; Glycoprotein; Hydrolase; Serine protease;
CC Signal; Zymogen.
FT SIGNAL 1 28 Potential.
FT PROPEP 29 32 By similarity.
FT CHAIN 33 260 Neuropsin.

```
FT ACT_SITE 73 73 Charge relay system (By similarity).
FT ACT_SITE 120 120 Charge relay system (By similarity).
FT ACT_SITE 212 212 Charge relay system (By similarity).
FT DISULFID 39 173 By similarity.
FT DISULFID 58 74 By similarity.
FT DISULFID 145 246 By similarity.
FT DISULFID 152 218 By similarity.
FT DISULFID 184 198 By similarity.
FT DISULFID 208 233 By similarity.
FT CARBOHYD 110 110 N-linked (GlcNAc...) (Potential).
FT VARSP LIC 23 23 A -> AACGSLDLLTKLYAENLPCVHLNPQWPSPSHCPRG
FT FTID=VSP 005401.
FT WRSNPLP PAA (in isoform 2).
SQ SEQUENCE 260 AA; 28048 MW; EF439E5B8C3E660 CRC64;

Query Match 39.9%; Score 641.5; DB 1; Length 260;
Best Local Similarity 42.6%; Pred No. 1.8e-50;
Matches 126; Conservative 43; Mismatches 80; Indels 47; Gaps 7;

QY 1 MATARP-----PMWVLCALITALLGVTEHVLANNVSCDHPSTNTPSGSNQDLGAGAGE 56
Db 1 MGRPRPRAAKTWFL-----LLLG-----GAWAGH 25

QY 57 DARSDSSRIINGSDCDMHTOPQOALLLRPNQLYCGAVLVHPQWLLTAACHCKKVPV 116
Db 26 SRAQED---KVLGGHECQPHSQPWQAA-LFQGGQLLGGVGVGGNWLTAACHCKPKYTV 81

QY 117 RLGHVSLSPVYESGQMFQGVKSIPIHPCGYSH---PGHSNDLMLIKLNRIRPTKDVREIN 173
Db 82 RLGDHSLQN-KDGPQEIPVVSIPHPCYNSDVEDHNDLMLQLRDQASLGSKVKRPFIS 140

QY 174 VSSHCPSAGTKCLVSGWGTTKSPQVHFPEKVLQCLNISVLQKRCEDAYPRQIDDTMFCAG 233
Db 141 LADHCTQPGQCTVSGWGTVTSPRENFPDTLNCAEVKIFPKCKCEDAYPGQITDGMWCAG 200

QY 234 DKAGRDSGQSDSGPVPVNCNGLQGLVSWGDIYPCARPMPGVYTNLCFTKWIQETI 289
Db 201 SSKGADTCQGDGGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKII 256
```

Search completed: July 4, 2005, 15:22:34
Job time : 73 secs